

Handbook of Theories of Social Psychology: Volume 1

Evolutionary Theory and Human Social Behavior

Contributors: Douglas T. Kenrick

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Evolutionary Theory and Human Social Behavior"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n2>

Print pages: 11-32

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n2>

[p. 11 ↓]

Chapter 1: Evolutionary Theory and Human Social Behavior

Douglas T. Kenrick, ed.

Abstract

From an evolutionary perspective, all the reactions people typically have to one another reflect the influence of psychological predispositions that helped our ancestors survive and reproduce. Any behavior can be understood at several levels of analysis, involving immediate triggers in the environment and the person's current biochemistry, developmental experiences, and evolutionary analyses of the adaptive function of certain choices over others. The evolutionary approach is thus not an alternative to other approaches discussed in this book. Instead, researchers adopting this perspective attempt to integrate research findings on ongoing social cognition and interpersonal relationships with theory and research from evolutionary biology, anthropology, and cognitive neuroscience. This chapter reviews research applying evolutionary models to: (1) biases in information processing (such as outgroup stereotypes), (2) the influence of affect and motivation on the decisions we make about other people, and (3) how simple, evolved decision-biases contribute to social dynamics and the emergence of culture. From this perspective, the mind is a coloring book, rather than either a blank slate or an unfolding blueprint. The approach has implications for important everyday behaviors and social problems, and the chapter concludes with a discussion of some of those implications for research on prejudice and economic decision-making.

Introduction

Evolutionary social psychology begins with a simple assumption, namely that people's interactions with one another are influenced by mental and emotional mechanisms shaped by natural selection. From this perspective, the way individual people think about one another, the way families, friends, and enemies feel about one another, and even the societies human beings construct, can be better understood by considering local social interactions in the broader context of other societies, and by considering all human societies in the still-broader context of other animal species. In a classic example of this approach, Charles Darwin suggested in 1872 that emotional expressions served an adaptive function – communicating one person's motivations and intentions to others (e.g., overt expressions of anger reduce the odds of [p. 12 ↓] costly physical conflict; smiles increase the odds of cooperation). Along with physical features such as upright stance and opposable thumbs, evolutionary theorists assume humans inherit brains equipped with mechanisms for managing our movement through the physical world (e.g., seeing in color, discriminating sugars from poisonous alkaloids) and the social world (e.g., speaking languages, bonding between mother and infant).

Although the central assumption at the heart of the evolutionary approach may not sound controversial, there has been debate about the extent to which evolved mechanisms are involved in human social behavior (see Alcock and Crawford, 2008; Kenrick, 2006; Tybur et al., 2008). Psychologists have wondered about questions such as: how can researchers sort out the influence of evolved mechanisms from effects of culture or learning? Consider the difficulties involved in trying to determine why women in North America are less likely to commit homicides than men. That difference could be a product of evolved mechanisms, or a product of the fact that American women grow up seeing movies and reading books in which men are depicted as more violent, or seeing boys, but not girls, encouraged to fight. If one assumes that culture, learning, and evolved mechanisms interact with one another, such questions become rather complex, requiring inputs from many different fields of research. In fact, modern evolutionary theory involves the integration of a broad network of ideas and evidence from different disciplines, including evolutionary biology, anthropology, and cognitive science.

Despite their complexity, the questions raised by an evolutionary perspective involve profoundly important issues about human nature and society. An evolutionary perspective has implications for every domain of human social behavior, from altruism, friendship, and love, to aggression, prejudice, and stereotyping. Evolutionary social psychology involves questions that are not only theoretically interesting, but also immensely important in a practical sense, with implications for law, business, medicine, and political science.

On Becoming an Evolutionary Social Psychologist

My initial interest in the evolutionary perspective could be attributed either to a broad-ranging curiosity about nature, or, less nobly, to an inability to decide what I wanted to be when I grew up. As an undergraduate I started off in biology, switched to psychology, and then considered switching to anthropology. I started graduate school in clinical psychology, but in my second year of graduate school I got to teach a lecture class in general psychology. I discovered that, if I were to become a research psychologist, I could study artistic creativity and mindless conformity, parental love and homicidal violence, sexual attraction and racial bigotry, or any other topic that involves thought, feeling, or behavior. This array of choices didn't require any premature foreclosures on my life options, so I decided to switch from clinical psychology to a research track in social psychology.

My conversion to an evolutionary perspective came two years later, as I was preparing for comprehensive examinations in social psychology. I should have been holed up in the library, reading all I could about experiments on dissonance theory, attributional processes, and objective self-awareness. But whenever I have a daunting amount of work to do, I develop intense interests in anything unrelated to the task at hand. In this spirit, I drifted into the campus bookstore and picked up a copy of *Primate Behavior and the Emergence of Human Culture*, by anthropologist Jane Lancaster (1975). This particular volume seemed comfortably outside the domain of experimental social psychology, so in my work-avoidance mode I felt compelled to buy it, bring it home, and read it. Lancaster's book indeed had little [p. 13 ↓] to do with the questions my social

psychology professors asked during my comprehensive examination. But I came to believe it raised many questions they should have asked.

Despite the fascinating range of research topics in social psychology, the scope of theory in the field was rather narrow at the time. Preparing for my comprehensive examination, I encountered a scattered disarray of unconnected small range theories, each independently advanced to explain a particular facet of social behavior. One mini-theory addressed frustration-induced aggression, another covered interpersonal attraction between people with similar attitudes, another addressed responses to one-sided versus two-sided arguments, and on and on. I missed the grand theories of personality and behavior I'd been exposed to during my clinical training, but when I mentioned this to the social faculty, they wagged their fingers and proudly informed me that social psychology was a "mini-theory" discipline.

Social psychologists at the time prided themselves not only on being theoretically constricted, but on being empirically narrow as well – studying anorexically thin slices of thought and behavior. Social psychologists in 1975 rejected the study of stable "traits" and focused mainly on people's thoughts and behavior in response to the "immediate situation," or at least those situations that could be captured within the half-hour duration of a typical psychology experiment. There were reasons for these strictures – experimental studies were designed to maximize control, and theoretical restraint was intended to cut down on rampant speculation about unobservable events inside the head or body – things scientists couldn't easily observe and count. But to a curious young student interested in the roots of human behavior, those constraints seemed excessive. I was not alone in being troubled by these limitations. Indeed, social psychology was undergoing an "identity crisis" at that time, with several of the field's leaders calling for an expansion of our theories and our research methods.

In this context, I took an almost guilty delight in glimpsing the very broad theoretical perspective suggested in Lancaster's book on primate social behavior. Instead of a narrow focus on one or another aspect of the social behavior of the members of our particular culture under specific laboratory conditions, Lancaster's evolutionary perspective offered a tantalizing suggestion – that we ought to erase the lines between psychology, biology, and anthropology, and consider how all these vast subjects fit together.

I began raving about Lancaster's book to anyone who would listen. Some of my fellow graduate students and faculty advisors just gave me an uncomfortable smile, as if I was earnestly explaining why I had just joined a cult. But a new assistant professor named Ed Sadalla had just picked up a copy of a recent book titled *Sociobiology* (Wilson, 1975), which considered common evolutionary principles underlying the social behavior of animals from ants to humans. Sadalla suggested that the evolutionary approach had great promise for generating testable hypotheses about human behavior, and he already had one he thought we should test. As part of a process Darwin called "sexual selection," females in many species choose males who have proven their dominance over other males, whereas males, with less to lose from an ill-chosen mating, tend to be less selective. Along with Beth Vershure, Sadalla and I began a series of studies suggesting analogous processes in humans. Although our results were clear and reliable across several experiments, it took us over a decade to publish them (Sadalla et al., 1987). As it turns out, we had unintentionally walked into an intellectual firestorm. There was an academic tumult surrounding sociobiology, which became the subject of a fascinating chapter in the history of science (Segerstrale, 2000).

The various controversies surrounding evolutionary social psychology were discouraging to many young researchers, but they have not been fruitless. Controversy often contributes to empirical and theoretical progress, as a theory's proponents search for [p. 14 ↓] new findings to address critics' skepticism. The evolutionary approach has generated many new findings and ideas, and the field's top journals have since published hundreds of social psychological studies testing evolutionarily informed hypotheses about the whole range of social behaviors, from altruism to xenophobia (e.g., Griskevicius et al., 2007; Navarette et al., 2009; Schaller and Murray, 2008). Many of today's prominent social psychologists, including several contributors to this volume, have incorporated evolutionary perspectives into their research (e.g., Brewer and Caporael, 2006; Cialdini et al., 1997; Deci and Ryan, 2000; Fiske et al. 2007; Higgins and Pittman, 2008; Huang and Bargh, 2008; Leary and Baumeister, 2000; Mikulincer and Shaver, 2006; Sedikides and Skowronski, 1997; Taylor et al., 2000; Van Knippenberg and Van Baaren, 2006; Van Vugt and Van Lange, 2006). At the same time, the approach continues to generate new empirical questions, and new theoretical puzzles to solve.

What is Evolutionary Social Psychology?

From an evolutionary perspective, all recurrent human social behaviors reflect the influence of physical and psychological predispositions that helped our ancestors survive and reproduce. This does not mean every individual social behavior is successful in promoting survival and reproduction, and it does not mean people (or other social animals) consciously think about survival and reproduction all, most, or even much of the time. It does imply that any social animal's brain is composed, in part, of mechanisms that helped its ancestors succeed in interactions with other members of its species. Thus, humans' reactions to other humans are presumed to *reflect the influence* of mechanisms shaped to solve the kinds of problems and opportunities our ancestors regularly encountered. I emphasize "reflect the influence" because an evolutionary approach does not imply that human behaviors are robotically determined by instinctive mechanisms over which we have no conscious control, or which are impervious to environmental inputs. People can and often do exercise control over powerful and fundamental emotional and motivational inclinations, including anger, fear, and sexual arousal. Furthermore, most mental mechanisms reflect the operation of flexible trade-offs, determined in interaction with current environmental conditions and past learning experiences (e.g., Ohman and Mineka, 2001). Although flexible, the influences of evolved predispositions (like hunger, thirst, sexual arousal, and anger) are nevertheless powerful vectors in our decision-making.

To understand the importance of evolved mental mechanisms, it helps to step back from our own species, and consider the interaction between other animals' bodily features and their environments. Killer whales, for example, though related to cows, would not do well with a cow's brain, since a killer whale's brain must control a body that tracks prey in the ocean rather than eating grass in a meadow. Likewise, bats, though also mammals, need brains designed to run tiny bodies that fly around catching insects at high speeds in the dark. Because all organisms' brains are composed of mechanisms evolved to deal with recurrent environmental threats and opportunities, evolutionary theorists ask: what are the implications of human evolutionary history (e.g., living in omnivorous and hierarchical primate groups populated by kin) for the design of the human mind? Evolutionary social psychologists focus on the subset of questions

dealing with recurrent social conditions of human life, and their hypotheses reflect anthropological data about social interactions common in societies around the world (e.g., close relationships with family, dominance hierarchies, long-term bonds between parents, common dangers from other groups competing for resources and territory, etc.), as well as general principles derived from placing humans [p. 15 ↓] in the context of other species confronting diverse adaptive problems (e.g., Allen-Arave et al., 2008; Lummaa, 2007).

What are the Roots of Evolutionary Social Psychology?

Around the time I was studying for my comprehensive examinations in 1975, there was an explosion of interest in, and controversy surrounding, the evolutionary approach to human behavior, centered around E.O. Wilson's book *Sociobiology*. But the perspective had been developing for some time, and represented the convergence of several streams of influence. A major contribution came from the field of ethology – the study of the behavior of animals in their natural habitats. In 1973, three European ethologists received the Nobel Prize for work on the adaptive significance of animal behaviors. One of them, Niko Tinbergen, experimentally studied how stickleback fish respond with complex behavioral displays to other sticklebacks (e.g., males demonstrate an aggressive display on seeing a red underbelly on another male, and the mechanism can be tricked by using “supernormal stimuli” – fishlike shapes with bright red paint on the underside). The second, Karl Von Frisch, demonstrated that bees engage in complex communications – informing other colony members about the location of nectar-bearing flowers. The third Nobel Prize-winning ethologist, Konrad Lorenz, conducted research on imprinting, the process by which young geese became attached to their mothers. Each of these lines of research demonstrated an interaction between an innate mechanism and inputs from the social environment.

Did the ethological work on other animals have implications for human social behavior? Many behavioral scientists at the time believed the answer was yes. Paralleling Tinbergen's findings on innate communication in bees, human language seemed a beautiful example of instinctive preparedness interacting flexibly with inputs from

the environment (Pinker, 1994). For example, comparative linguists had uncovered evidence that languages the world over have a similar underlying grammatical structure, developmental psychologists found that children make similar linguistic mistakes regardless of the language they are learning, and the languages spoken by adults everywhere were found to be equally complex, regardless of the speaker's education level. Paralleling Lorenz's work on animal imprinting, developmental psychologists had found that young children go through predictable phases in their relationships with parents and strangers (such as a fear of strangers that peaks around nine months), and that infants engage in complex nonverbal mimicry of their mothers. Paralleling Tinbergen's research on fixed action patterns in sticklebacks, Irenaus Eibl-Eibesfeldt (1975) found a sequence of nonverbal flirtation gestures in societies around the world, and other research revealed that children born blind and deaf nevertheless demonstrated appropriate facial expressions in appropriate contexts (e.g., smiling when tickled) (Eibl-Eibesfeldt, 1973).

I had personally observed the power of innate mechanisms on animal social behavior as a child, when I raised tropical fish for a hobby. Following instructions in a tropical fish manual, I placed a male *Betta splendens* (or Siamese fighting fish) in a tank, fed him an abundant supply of live food, and raised the water's temperature and acidity (thereby simulating conditions during breeding season in Southeast Asia, from whence his ancestors came). He responded by building a nest of bubbles on the water's surface, exactly as my manual had predicted. When the nest reached the right size, I introduced a female *Betta* who had been exposed to similar conditions and developed a bulge in her belly, indicating a supply of eggs. Upon seeing the female, the male puffed out his colorful fins, and began swimming around her in a mating dance. She responded with movements indicating her readiness to mate, and he then wrapped his [p. 16 ↓] body around hers, triggering her release of several eggs and his release of sperm. He scooped up the eggs in his mouth, placed them in the bubble nest, and repeated the sequence several times until her supply of eggs was depleted. He then chased her away and began a period of guarding the eggs and later the emerging tadpole-like fry. This sequence was complex, but none of it was learned. I know this because I watched one of his offspring perform the exact same ritual with another female, despite the fact that he had been raised in isolation (which is common with male Siamese fighting fish, who are notoriously pugnacious and likely to attack other fish).

The ethologists' research made it clear that much so-called instinctive behavior (such as imprinting) involved what Ernst Mayr called "open instincts" – innate proclivities requiring environmental inputs to be fully operative. Although a fixed sequence of movements is fine for mating rituals, some innate predispositions require flexibility to be useful. For example, animals need a rapid response system for avoiding other dangerous animals and poisonous foods, but exactly which other species in the local environs pose threats is often quite variable over the range into which they might be born (consider young English sparrows who, like humans, may be born in deserts, forests, farmlands, or cities on several continents). Hence, those avoidance systems need to be calibrated to reflect local threats (Ohman and Mineka, 2001).

Psychologists studying learning in the laboratory had presumed mammals possessed only a very few innate drives (such as hunger and thirst, often referred to as "primary drives"), and that experiences after birth led to the development of "secondary drives" (desires for other stimuli associated with satisfaction of hunger and thirst). The development of secondary drives was believed to depend on two simple forms of learning – classical and operant conditioning, the rules of which presumably applied similarly to many kinds of learning across many species (Skinner, 1953). This view had the advantage of being parsimonious, meaning it could explain much with few assumptions. However, several behavioral psychologists began to uncover findings challenging this view. The rules of conditioning changed depending on what was being learned and which species did the learning (Seligman and Hager, 1972). Consider the rules involved in learning to avoid poisonous foods. Some learning requires instantaneous feedback (e.g., a jolt of pain immediately after touching a hot stove facilitates learning not to touch it again), but people and other animals learn to avoid foods that made them sick many hours after the food was eaten. Learning food aversion is also unlike many other types of learning in that it requires only one trial and is difficult to extinguish. Furthermore, the types of stimuli that get conditioned to nausea vary in ways consistent with the organism's evolutionary history and typical ecology. For example, rats, which have poor vision and rely on taste and smell to find food at night, easily condition aversions to novel tastes, but not to novel visual stimuli (Garcia and Koelling, 1966). Quail, on the other hand, have excellent vision and rely on visual cues in food choice, and show the opposite bias, conditioning easily to visual cues but not to taste (Wilcoxon et al., 1971).

Although not many social psychologists were radical behaviorists, most shared the assumption that social behaviors were products of individual learning experiences and not innate mechanisms; however, two other influences at the time began to challenge that simple empiricist world view. Behavioral geneticists were uncovering evidence that complex human behaviors could be passed on genetically. For example, twins raised separately shared many adult personality traits, including rare forms of psychopathology, whereas adopted children shared surprisingly few such traits with the families in which they were raised (Rowe, 1994). At the same time, advances in cognitive science, including the study of artificial intelligence, were beginning to suggest that complex cognitive processing could be programmed [p. 17 ↓] from very simple “on-off” components (Gardner, 1985).

At the time, social psychologists were, like ethologists, becoming more open to studying behavior in naturalistic contexts. In particular, researchers interested in relationships were arguing that processes unfolding over months and years of social interactions might not emerge in a half hour between strangers in laboratory experiments (e.g., Murstein, 1981).

Multiple Levels of Causal Analysis

Even in the 1970s, most behavioral scientists accepted natural selection as an explanation of animal morphology (why giraffes have long necks and porpoises have fins, for example). They also accepted that human morphology, including upright stance, opposable thumbs, and large brain size, were products of natural selection; likewise for the behavior of other animals, such as peacocks' mating displays. But back then, it was less clear how an understanding of natural selection applied to human thought and behavior, and if so, how that understanding would translate into hypotheses of the sort typically considered by psychologists. My friend Jim Sherman, for example, had taken E.O. Wilson's biology course when he was a young student at Harvard. Jim did not doubt the theory of natural selection, but when we discussed these issues in the 1980s, he questioned whether evolutionary theory could yield testable implications for social cognition. Jim suggested I consider switching my field of interest to anthropology or biology.

Other behavioral scientists at the time and since have questioned how one can ever study the evolution of behavioral inclinations. Unlike bones, behaviors generally do not leave fossil evidence. Although anthropologists are clever in deriving evidence about ancestral behavior, many interesting behaviors, such as flirtatious gestures or gossip-filled conversations, leave no traces anyone has yet deciphered using archaeological or geological methods. But even if there were major advances in such inferential techniques, most evolutionary social psychologists would not buy a pith helmet and rush off to the Olduvai Gorge. Why? Because an evolutionary perspective on psychology does not typically involve a search for historical roots of behaviors – for when and how particular behavioral mechanisms evolved. This is also true of animal ethologists studying insects, birds, and fish; they are not typically concerned with the evolutionary history of the inclinations they study either. To understand why not, it is important to appreciate a distinction biologists frequently make between explanations at different levels of analysis.

Historical controversies have been fueled by failures to distinguish between different levels of causation within the field of biology. Consequently, biological theorists have stressed the importance of differentiating between causal explanations of behavior involving evolutionary history, adaptive function, ontogenetic development, and proximate determinants (e.g., Sherman, 1988). For example, consider the question of why human mothers nurse their offspring. This can be addressed at four different levels of analysis:

- 1. *Historical* explanations consider *ancestral roots* of behaviors. Researchers adopting a historical perspective might consider human nursing in the comparative context of other animal species. Unlike human language, nursing capacity does not pose much of a historical puzzle, since all primate females nurse their offspring, as do all mammals.
- 2. *Functional* explanations, on the other hand, are concerned with ultimate *adaptive purposes* of behaviors. A functional explanation might suggest mothers nurse offspring because it increases offspring survival rates.
- 3. *Developmental* explanations are concerned with *lifespan-specific inputs* that sensitize the organism to particular cues. A developmental explanation might suggest that mothers nurse offspring because pregnancy and

childbirth trigger puberty-dependent shifts in hormones and milk production in mammalian females.

- 4. *Proximate* explanations focus on *immediate triggers* for particular behaviors. A proximate explanation might suggest that nursing occurs because an infant has begun suckling on the female's nipple, which leads to immediate hormonal changes inside the mother, which stimulates milk release.

Sometimes there are obvious connections between the different levels of analysis. In the case of nursing, developmental changes in lactation capacity accompany other changes during pregnancy, and the infant, who receives obvious functional benefits from nursing, triggers the proximate release of milk. But connections between different levels of analysis are not always obvious. Consider why birds migrate each year. A proximate explanation is that birds migrate because days are getting shorter – the immediate cue triggering migration. The ultimate reason for such migration, however, is survival and reproduction – the distribution of desirable food and mating sites varies seasonally. There are two key implications here: (1) animals, including humans, need not be consciously aware of the ultimate functions of their behaviors; and (2) the connection between long-term goals and immediate goals is often indirect and not obvious.

A given researcher is usually concerned with one or two levels of analysis. However, an explanation at one level of analysis must be compatible with explanations at other levels. Positing a proximate or developmental mechanism that reliably leads people to make functionally maladaptive decisions (such as Freud's death instinct) is problematic. Evolutionary psychologists typically advance hypotheses about links between proximate mechanisms and adaptive function, not about the historical roots of the mechanism. In deriving those hypotheses, however, psychologists adopting an evolutionary perspective attempt to take into account pertinent findings from evolutionary biology and/or anthropology. Psychologists can derive hypotheses about proximate causes and development without thinking in evolutionary terms, but disregarding evidence and theory derived from research on other cultures and other species can lead to hypotheses incompatible with other levels of analysis. For example, psychologists during the last century often assumed most sex differences in social behavior (such as differences in violent aggression) were products of "American

culture,” unaware that similar differences were found in other cultures, and even other species. Indeed, the study of sex differences in social behavior led to some of the first critical findings in evolutionary social psychology, as I will describe below.

Critical Tests versus Nomological Networks

Could an evolutionary perspective provide new and testable hypotheses about human social behavior? A natural place to begin looking was by considering sex differences in mating-related behaviors. Sex differences in mating behavior are found throughout the animal kingdom, and comparative biologists have made progress understanding not only why those differences exist, but also why they are sometimes relatively larger or smaller. Darwin himself wondered why there are such dramatic sex differences in some species (such as peacocks and peahens), and why some features (such as peacock's brilliant and attention-getting feathers) involve features likely to decrease survival. He coined the term *sexual selection* to refer to the evolution of features useful not for survival, but for either competing with members of the same sex or attracting members of the other sex. A century later, Trivers (1972) connected sexual selection to *differential parental investment*. Within any given species, there is often an initial discrepancy in minimum resources or effort one sex is required to invest in the offspring. In mammals, for example, females must carry [p. 19 ↓] offspring inside their bodies, and later nurse them for weeks, months, or years. On the other hand, the minimum obligatory investment by mammalian males is a few moments of courtship and the energy required to copulate. Because mammalian females are limited to fewer offspring than a male could have (if he could attract and inseminate multiple females), females choose more carefully, picking males demonstrating either superiority to other males, or willingness to invest effort and resources in the female and their offspring.

The research Sadalla, Vershure, and I started in 1976 was designed to test a simple derivation from the literature on sexual selection and differential parental investment – that women, like females in other mammalian species, would be attracted to males giving behavioral indications of dominance relative to other males. We presumed that men, like males in other mammalian species, would be relatively less attentive to female dominance. In a series of experiments manipulating social dominance in various ways, we found support for that hypothesis. Early reviewers, however, were not convinced

these findings clearly favored an evolutionary explanation over what they considered a “more straightforward” explanation based on learned social norms. According to this alternative, people are attracted to members of the other sex who conform to societal expectations, which differ for men and women. Reviewers argued that American men were expected to act dominant, and women to act submissive. We saw problems with this alternative. First, it did not explain why men in our studies were not attracted to submissive females (who were, after all, conforming to the presumed societal expectations of the time). Second, the alternative explanation tacitly assumed the cultural norms were arbitrary, without considering why it was more socially normative for men to act dominant in the first place. Third, although the reviewers expressed confidence that sex differences in dominance were products of American culture, they did not present evidence of cultures in which females were more dominant than males, and in which men, but not women, preferred dominant partners. As later research indicated, women all around the world care more about social dominance and status in their mates than men do (Buss, 1989). Fourth, the alternative could not explain why the sex difference among North Americans was also found in many other animal species. Finally, it could not account for findings that the hormone testosterone, found in much larger quantities in male mammals, was closely linked to dominance behaviors and mating displays. Our findings did not, by themselves, rule out the possibility that North American society had constructed this sex difference arbitrarily, and that the apparent links with other species were mere coincidence. However, our findings fit into a broad *nomological network* of research findings from several disciplines. Indeed, to appreciate the strength of evolutionary explanations, it is often necessary to appreciate how a particular finding fits with many other findings.

Our findings on dominance and sexual attraction came from laboratory experiments involving North American college students, and could not, in themselves, rule in favor of an evolutionary explanation over a cultural or learning alternative. In later research, Rich Keefe and I adopted a different approach, examining archival data on sex differences from many different cultures. Keefe and I were interested in another sex difference – the common tendency for younger women to pair up with relatively older men (Kenrick and Keefe, 1992). Several social scientists had conducted analyses of singles' advertisements, and observed that men typically advertised for a partner a couple of years younger, and women typically wanted a partner slightly older than themselves.

Although this violated the general tendency to prefer partners as similar as possible, it was typically attributed to the influence of sex-discrepant norms in American society (e.g., Cameron et al., 1977).

[p. 20 ↓] Keefe and I suspected the sex difference in age preferences might be better explained in terms of ideas from a branch of evolutionary theory called *life history theory*. Life history theorists explore how each animal's life cycle – from conception to death – is shaped by natural selection to facilitate reproductive success (Stearns et al., 2008). A life history is a genetically organized developmental plan – a set of general strategies and specific tactics by which organisms allocate energy to survival, growth, and reproduction. Different species have very different life histories; some (such as salmon) reproduce in a single burst of effort, others (such as humans) reproduce many times over their life span. Some animals begin reproducing shortly after birth; others, such as humans and elephants, wait over a decade to begin reproducing. Within any species, there are often differences in life history, some linked to the differences in parental investment we discussed earlier. In the case of humans, females typically reach sexual maturity a year or two earlier than males, a difference common in species in which larger males compete more successfully for mates (Geary, 1998). Human females, unlike most other mammals, go through age-linked physiological changes terminating their reproductive capacities. Menopause may have evolved because ancestral women lived long lives, and typically invested heavily in grandchildren, balancing costs of direct reproduction against benefits of indirect kin care.

Unlike women, men do not directly invest bodily resources in offspring. They do, however, frequently invest indirect resources, providing food, care, and protection. Anthropological studies of hunter-gatherers suggest that men become increasingly skilled at providing resources as they age (Kaplan et al., 2000). Even in modern societies, men's social status and resources typically increase for several decades after reaching sexual maturity. Males do not go through menopause and are capable of fathering children into their 80s. All of these considerations led Keefe and I to guess that sex differences in age preference were better explained in terms of human life history than American cultural norms. On the life history view, females choose older males who traditionally had more access to resources; males prefer females with features indicating peak fertility (between late teens and early thirties). If the evolutionary life-history account were true, then preferred age differences between partners ought not to

be constant, but to change depending on each partner's age. In particular, preferences for younger females would not be pronounced among young men, but would get stronger as men aged (and women their age had fewer remaining years of fertility). Young men, although typically most sex-typed, would violate this presumed norm by expressing attraction to women slightly older than themselves. Those predictions were confirmed in several studies, raising difficulty for explanations based on sex-typed norms (Kenrick and Keefe, 1992; Kenrick et al., 1996a).

More critically, the life history perspective suggested that this sex difference would be found across human societies, and not be in any way associated with North American cultural norms. Indeed, we found the sex difference in age preferences across societies ranging from Germany to India (despite many other differences in marital and mating customs). We also found it across historical periods, on a small remote island in the Philippines in 1913 and in the Netherlands during the 1600s and 1700s. Consistently, the United Nations data on marriages indicated that young men around the world are several times less likely to marry than are young women, whereas older men are much more likely to marry than are older women (Kenrick and Keefe, 1992). The size of the sex difference varies somewhat across societies (Eagly and Wood, 1999; Kenrick and Keefe, 1992), but contrary to social norms explanations, the difference is actually least pronounced in North America and Europe in comparison to less Westernized societies. As I'll discuss later, the cultural variations are themselves likely to be [p. 21 ↓] linked to biological factors (Kenrick et al., submitted).

Other researchers have found various universals in human social behavior. For example, romantic love, once believed to be a cultural invention of medieval European societies, is found all around the world (Jankowiak and Fisher, 1992). Sex differences in violent crime, again often attributed to American cultural norms, are found around the world and over historical periods (Daly and Wilson, 1988). Buss (1989) and Schmitt (2003) and their colleagues have also found numerous universals in human mating arrangements.

Cross-Fertilization versus Artificial Insemination: Bridging Social Psychology and Evolutionary Theory

Much of the original debate involved attempts to uncover findings that fit better with evolutionary models than with traditional social science models, to support the premise that pursuing an evolutionary perspective was not a waste of time. So far, the story sounds more like artificial insemination than cross-fertilization, with evolutionary biology providing important ideas and psychology absorbing them. But social psychologists have many ideas, findings, and research tools that have contributed to a better understanding of the links between evolution and social behavior.

Consider gender differences in mating-relevant behaviors. Evolutionary theorists initially emphasized these differences, following considerations of life history and parental investment. Although sex differences are often large, however, they are sometimes small. Social psychological research tools and theory allowed us to address why. Recall that social psychologists in the 1970s had begun arguing that laboratory experiments might not be ideal for studying friendships and romantic relationships. Those researchers demonstrated that what people find rewarding at a relationship's beginning is not necessarily what they find rewarding in long-term relationships or marriages (Murstein, 1981). This insight, when integrated with the theory of differential parental investment, was important to understanding variations in sex differences and similarities in mate choice. In the initial phases of romantic relationships, men have invested very little; yet if a woman were to become pregnant, she would invest a great deal in any resulting offspring. Over time, men invest more resources and effort in romantic relationship, as in a marriage. Across species, when males invest little in offspring, only females are choosy about mating; as males invest more, they become increasingly selective. Indeed, in some species of birds (such as phalaropes) and fish (such as seahorses), males invest more than females in the offspring, and everything changes (males become more choosy about mating; females become more competitive with one another to obtain males). Many species, including humans, have more than one mating option (Gangestad et al., 2007). As in most other mammals, it is possible

for human males to make little investment, in which case, females ought to select males based on evidence of “good genes” relative to other males (using cues such as appearance and behavior, as peahens do in distinguishing between peacocks). When, alternatively, human mating involves high levels of male commitment, men ought to be more selective, as in other species with high male parental investment.

We tested the hypothesis that human mating preferences show contingent changes with level of commitment in several ways. In some studies, we simply asked men and women to specify their “minimum standards” for various characteristics in single dates, sexual partners, steady partners, or marriage partners. Men and women specified similarly high standards for marriage partners, but very different standards for sexual partners (Kenrick et al., 1990). Indeed, in considering one-night stands, men were [p. 22 ↓] willing to consider partners below their normal standards for a date (Kenrick et al., 1993).

Although both sexes became more demanding for long-term partners, there were still differences in the particular traits they most valued. In particular, women placed much more emphasis on status and wealth in long-term partners; men placed somewhat more importance on physical attractiveness. Other evolutionarily oriented researchers have found evidence that physical attractiveness is linked to cues signaling a woman's fertility, whereas status and wealth are linked to men's ability to generate resources. In later research, my former student Norm Li, who had been trained in economics before switching to social psychology, devised a method for distinguishing between traits people regard as “luxuries” versus “necessities.” Previous researchers had often simply asked people what they most preferred in mates, without considering what those people might actually be able to “afford.” If you simply ask someone what they prefer in a mate, both men and women are attracted to someone who is as physically attractive as a movie star, as creative and funny as a comedian, as warm as a loving grandmother, and as wealthy as Bill Gates. Indeed, when Norm Li gave participants a high budget of “mating dollars,” men and women differed only slightly in their criteria for mates. But what if they are given a more realistic budget, one like most mortal humans actually confront when choosing a mate? With a limited budget, when choosing a lot of x meant compromising on y , the sexes expressed very different preferences. Women regarded social status as a “necessity,” and were willing to compromise on physical attractiveness. Men regarded good looks as a necessity, and were willing to disregard

status and other characteristics that would otherwise be desirable (Li and Kenrick, 2006; Li et al., 2002).

In recent years, there have been a number of other findings especially difficult to explain except in light of an evolutionary perspective. For example, several recent studies suggest that when females are ovulating (and most likely to become pregnant), they are more attracted to men showing indications of “good genes,” including high levels of masculinity and symmetry (e.g., Penton-Voak et al., 2003). Ovulating women are also more interested in extra-pair sexual relations, particularly with men more attractive than their long-term partners (Pillsworth and Haselton, 2006). Most interestingly, ovulating women are more likely to express attraction toward the smell of highly masculine and symmetrical men, based merely on exposure to t-shirts worn by these men (Gangestad and Thornhill, 1998).

Why the greater emphasis on physical attractiveness for extra-pair sexual liaisons than for long-term relationships? Presumably, symmetry and high masculinity in men, like colorful and symmetrical displays in peacocks, indicates the possession of genetic traits well-suited to survival. If a female has a casual sexual liaison, it could result in the transmission of the attractive male's beneficial genes to offspring, even though the male will not be around to provide resources. Hence, females are presumed to face some degree of trade-off between choosing a male who will stay around and provide resources and one who is highly attractive to other females. Concealed extra-pair liaisons with highly attractive males during the period of maximal fertility are presumably a way of attempting to have it both ways. Presumably, such affairs are temporally limited and hidden because they result in potentially high cost (loss of the long-term partner) if discovered. None of this is consciously mediated, and the cyclic effects are not found for women on birth-control pills (which change normal hormonal patterns).

Biases in Information Processing: The Old Look

The evolutionary perspective was emerging during the period when the social cognitive [p. 23 ↓] perspective came to predominate in social psychology. When Jim Sherman

and I debated the relevance of an evolutionary perspective during the early 1980s, we perhaps viewed the two perspectives as alternative paradigms for our field. Many researchers have since realized that cognitive and evolutionary perspectives are not only compatible, but that a full understanding of human nature requires an integration of the two perspectives (e.g., Cosmides and Tooby, 1992; Haselton and Nettle, 2006; Kenrick, 1994, 2001).

An appealing feature of cognitive methods is that they can help bypass social desirability biases. If a research participant fills out a questionnaire regarding mating behaviors or preferences, he or she may strategically choose a socially appropriate response. But if challenged to identify a particular type of face in a crowd as rapidly as possible, or to say whether he or she remembers seeing a face previously flashed on the screen, or to judge the frequency of faces in a particular category, the bias is to try to get the right answer. Along with several students and colleagues, I have been involved in a program of studies examining the implications of evolutionary ideas for processes of attention, encoding, and retrieval (e.g., Ackerman et al., 2009; Kenrick et al., 1994, 2007; Roese et al., 2006). We find that people selectively attend to other people with functionally relevant characteristics, and are better at encoding and remembering those individuals.

Harking back to Darwin's evolutionary view of emotional expressions, another person's anger is a very functionally relevant stimulus, foreboding potential harm. If that other person is a male, the dangers are substantially higher: males commit approximately 90 percent of homicides, and even when females do commit murder, it is often a self-protective response to male harassment (Daly and Wilson, 1988). Males are also more likely to do physical damage if they hit or strike someone (Archer, 2000). Women, on the other hand, are more likely to afford cooperative opportunities to others (e.g., Taylor et al., 2000). Consistent with functionally derived hypotheses about emotion recognition, we found that participants not only identify whether a face is angry or happy in a fraction of a second, but are significantly quicker and more accurate at recognizing anger on a male than a female face, and happiness on a female face (Becker et al., 2007). Interestingly, the reverse is also true, people can identify a face as male or female in a fraction of a second, but are significantly faster at identifying a face as male if it is angry and as female if it is happy. To deal with the possible confound that men might be better at expressing anger and women better at expressing happiness, we

showed participants computer-generated faces made to look like males or females, and found the same effects. If a face was completely androgynous, but made to look slightly angry, participants judged it to be a man. Initially, some reviewers were confident these effects were products of cultural stereotypes associated with masculinity and femininity. We were able to design an experiment to pit the alternative explanations against one another. When we created androgynous faces and gave them masculine cultural cues (a suit and tie), they were judged as highly masculine, but not angry. When the identical faces bore feminine cultural cues (a blouse, earrings, and necklace), they were judged as feminine, but not happy. One critical cue was the masculine eyebrow ridge (which is lowered in males and in anger). When the androgynous face was given a lower eyebrow ridge, it was judged as significantly more angry, but only slightly more masculine (not as much as when it wore a suit and tie).

In another series of studies linking attention, encoding, and memory, participants were shown crowds of faces containing attractive and average-looking people of both sexes, and later asked to remember which faces they saw. Using an eyetracker, we found that both sexes looked relatively more at attractive women. When asked to judge the frequency of attractive women in a crowd, both sexes overestimated the number if the [p. 24 ↓] crowds were presented rapidly, suggesting that attractive women are more immediately cognitively accessible. Consistently, people of both sexes are good at remembering whether or not they have seen a particular attractive woman before (Maner et al., 2003). For attractive male targets, on the other hand, there is an interesting disjunction between different levels of processing. In several studies, we found that women (but not men) attend more to good-looking men, but then neither remember them well, nor overestimate their frequency in previously presented crowds. These findings make sense in terms of male and female mating strategies discussed earlier. For men, strange attractive women are mating opportunities. For women, who are typically more inclined toward long-term relationships with familiar men, and who typically do not make advances toward strangers, there is less inclination to devote processing time to strange attractive men.

Another study in this program used functional logic to reconsider a bias in memory linked to “outgroup homogeneity” and the failure to remember individual members of other racial groups. In functional terms, it makes sense to make finer discriminations between members of one’s own group (with whom one has more frequent and more

cooperative interactions). However, there are also times when it makes functional sense to pay close attention to, and remember the identity of, members of other groups – when those individuals might pose threats to one's welfare (members of outgroups often have less motivation to inhibit aggressive inclinations than members of one's own group). Consistent with that logic, we found the standard outgroup homogeneity effect when white participants were asked to remember faces of emotionally neutral black men, but a tendency to remember black men especially well if they were angry (Ackerman et al., 2006). When cognitive resources were limited, in fact, the outgroup homogeneity effect was reversed for angry faces, with white subjects remembering angry black men better than white men.

Affect, Cognition, and the Modular Mind

Evolutionarily informed research on learning and memory suggests that brains are composed of a number of specialized “domain-specific” mechanisms. For example, birds use different memory systems and different rules to remember species song, tastes of poisonous food, and locations of food caches (Sherry and Schacter, 1987). Many birds learn their species song during an early critical period, then reproduce it perfectly during the next breeding season, without ever having practiced it. On the other hand, birds learn the characteristics of poisonous foods in a single trial during any time of life. Following still different rules, food locations are learned, updated, and erased on a daily basis. Using the same decision rules for each of these problems would be highly inefficient, and different memory systems in birds are anatomically distinct. Likewise, humans inherit different memory systems to deal with different, sometimes conceptually incompatible tasks, including learning language, learning to avoid poisonous foods, and remembering other people's faces.

The notion of domain specificity has profound implications for studying the links between motivation and social cognition (Kenrick and Shiota, 2008; Kenrick et al., 2003; Neuberg et al., 2010). This approach assumes cognitive processes differ qualitatively depending on which fundamental social motive is currently active (see [Table 1.1](#)). Thinking in these terms leads to a number of hypotheses about how attention, memory, and downstream social interactions change in functionally specific ways as a function of activating motives such as self-protection, mating, or affiliation.

Social problem domain	Example of decision bias
Coalition formation	Cooperate with those who (a) share your genes, (b) have cooperated with you in past
Status seeking	Males will take more risks to gain and maintain status
Self-protection	Potential threats or costs trigger reciprocal aggressive behavior, particularly among non-kin
Mate choice	Males, compared to females, generally more inclined toward unrestricted sexual strategy
Mating relationship maintenance	Breaking a bond likely for: <ul style="list-style-type: none"><li data-bbox="894 967 1432 1115">• (a). males when a mate is sexually unfaithful or when physically attractive alternatives are available<li data-bbox="894 1115 1432 1262">• (b). females when a mate compromises resources or when high status alternatives are available
Parental care	Familial provision of resources and care will follow the order: <ul style="list-style-type: none"><li data-bbox="894 1368 1253 1410">• (a). self > siblings<li data-bbox="894 1410 1253 1474">• (b). own offspring > stepchildren

For example, in line with our earlier discussion of sex differences in parental investment and sexual selection, activating mating goals leads men to erroneously perceive sexual arousal (but not other emotional states) in photographs of neutrally expressive [p. 25 ↓] attractive women. No such effect emerges among female perceivers (Maner et al., 2005). Activating self-protection goals, on the other hand, triggers particular outgroup stereotypes connoting threat or danger, but not equally negative, but threat-irrelevant

stereotypes. For example, white North Americans who are in a dark room, or have just seen a fear-inducing movie, view black men, but not whites or black women, as dangerous or angry (e.g., Maner et al., 2005; Schaller et al., 2003).

Fundamental goals that seem, on the surface, specific to a narrow kind of behavior (e.g., mating) may have functional consequences spanning a broad range of behavior. For example, activating self-protective goals enhances conformity among both men and women; activating mating goals enhances conformity among women but leads to counter-conforming behavior among men (Griskevicius et al., 2006b). Another series of studies finds that activating short-term mating goals leads men, but not women, to be more creative (Griskevicius et al., 2006a). These findings all fit with predictions about differential consequences of different social behaviors depending on the functional goal accessible at the time.

Interesting questions arise from considering implications of this approach for other views of motivation (Kenrick, Griskevicius, et al., 2010). For example, regulatory focus theory (Higgins, 1998) links prevention-based thoughts and behaviors to the goal of security and promotion-based thoughts and behaviors to nurturance. This would suggest that the domain of self-protection is more closely related to prevention goals, and coalition and parental care to promotion goals. At the same time, each of the domains described in [Table 1.1](#) probably involves, to some degree, both promotion and prevention subgoals (finding new friends versus avoiding loss of existing friends, encouraging bonds with a mating partner versus avoiding infidelity, etc.). There are potentially productive empirical questions involving the extent to which similar or different types of cognitive biases are associated with different types of promotion and prevention in particular domains.

Social Dynamics and the Emergence of Culture: The Mind as a Coloring Book

Evolutionary social psychology is one component of a broader interdisciplinary movement toward integrative science (Kenrick, 2001; Wilson, 1998). Dynamical systems theory is another attempt to combine ideas from biology, mathematics,

computer science, and cognitive psychology (Latané, 1996; [p. 26 ↓] Nowak and Vallacher, 1998). Two fascinating conclusions have emerged from research on complex systems ranging from ant colonies to world economies. Dazzling fractal images represent one lesson of dynamical systems theory: awe-inspiring complexity can emerge from a few variables interacting according to a few simple rules. Such emergent processes can be intellectually overwhelming, but there is another side to the story. Complex systems at every level from molecules to genes to neurons to ecosystems often reveal *self-organization* – order emerging out of initial disorder.

Exactly which patterns of self-organization arise in complex systems depends on simple decision rules at the individual level. Combined with ideas from evolutionary modularity (e.g., biases in [Table 1.1](#)), this helps explain why different patterns of social organization arise in different domains of social life, such as the hierarchies associated with status, dyadic pairings associated with mating, and large aggregations associated with intergroup conflict (Kenrick et al., 2003). This work has important implications for another area of modern social psychology: the emergence of cultural norms (Harton and Bourgeois, 2004; Kenrick et al., *in press, b*). A dynamical evolutionary approach helps us understand how cultural norms encompass underlying universal biases as well as diverse cultural norms. For example, most human societies involve more marriages between older men and younger women than the converse, as discussed. There are, however, a few societies, such as Tiwi aborigines of Australia, in which older women marry younger men. On closer examination, those societies do not involve exceptions to underlying human nature, but a new dynamic emerging from multiple evolved biases. Men in Tiwi society are in fact attracted to younger women, and Tiwi women are attracted to older high status men. The Tiwi arrangement is an accommodation to another general feature of human social life – male competition for status and emergent patterns of patriarchy. In Tiwi society, older men have all the power, and, since they are polygynous, they betroth their daughters to other patriarchs, who reciprocate. This leaves the younger men out of luck. Another norm in Tiwi society, however, is that all women – of all ages – must be married. When an older woman becomes widowed, she must remarry, but the older patriarchs with younger wives do not wish to marry her. Instead, she marries a younger man, who thereby does a favor for her older male relatives, and puts himself into the game of acquiring younger wives in the future. These arrangements suggest that the mind-culture interaction is neither captured by

the “blank slate” metaphor or by the metaphor of a predetermined blueprint. Instead, we can fruitfully conceptualize the interaction using the metaphor of a coloring book, which has some pre-drawn lines, but a great deal of flexibility within which local colors can be recombined in diverse ways.

Conclusion: Applying Evolutionary Principles to Social Life

There is now abundant evidence supporting the premise that human social behaviors reflect the influence of physical and psychological predispositions that helped our ancestors survive and reproduce. Social psychology is increasingly moving to the center of the interdisciplinary synthesis constituting evolutionary psychology. Researchers are realizing that understanding human nature also has implications for social problems at all levels – from family conflicts and neighborhood crime to overpopulation and international conflict (Crawford and Salmon, 2004; Penn, 2003). Social psychologists have begun applying functional analyses to a wide range of social issues and interpersonal problems, including prejudice and intergroup conflict (Kurzban et al., 2001; Navarette et al., 2009), sexual harassment (Haselton and Buss, 2000; Kenrick et al., 1996b), homicide (Daly and Wilson, 1988; Kenrick and Sheets, [p. 27 ↓] 1994), and political conflict (Sidanius and Pratto, 1999; Weeden et al., 2008). An evolutionary approach also has implications for understanding clinical disorder (Keller and Nesse, 2006) and positive psychology (Buss, 2000; Kenrick et al., submitted).

Consider for example, the topic of prejudice. Although social psychologists typically conceived of prejudice as “negative affect,” Cottrell and Neuberg (2005) fruitfully expanded this conception by applying notions of functional domain specificity. They presented evidence that different social groups (e.g., African American males, Asian females, gay men, religious fundamentalists) often elicited qualitatively different emotional reactions linked to the particular functional threats posed by those groups. Other research suggests stereotypical and prejudicial responses to particular groups vary in functional ways with the nature of the threat and the functional context. For example, Schaller et al. (2003) found that being in a dark room specifically enhanced perceived physical threats from black and Arab men, but did not enhance other

stereotypical negative characteristics. And although prejudicial responses to these groups appear to be linked to fear, prejudicial responses to other groups (such as people from little-known foreign countries) are more associated with concerns about disease (Faulkner et al., 2004). As these authors point out, attempts to reduce prejudice and discrimination will be more successful to the extent that they target the particular functional motivations and the particular threats associated with particular groups.

An exciting recent development involves bridges between evolutionary psychology and economics (e.g., Ermer et al., 2007; Kenrick et al., 2008, in press, a; Wilson and Daly, 2003). As one example, a classic finding from behavioral economics involves “loss aversion” (Tversky and Kahneman, 1991). People are willing to pay more to ensure against the loss of \$100 they already have in their wallets than to gain an additional \$100. In one recent series of studies, we find that activating mating or status motives erases loss aversion (so that participants now value a given-sized gain as much as, or more than, an equal loss). In line with findings reviewed earlier that mating strategy involves more competitive showing off for males than for females, this mating-based reversal of loss aversion was found only for males (Li and Kenrick, submitted).

Thus, the evolutionary perspective scores highly on two key criteria for powerful scientific theories: comprehensiveness and heuristic potential. I have only touched on a somewhat idiosyncratic sampling of the broad range of hypotheses generated by this perspective (see Crawford and Krebs, 2008; Dunbar and Barrett, 2007, and Schaller et al., 2006 for a diverse sampling of other perspectives). Studying social psychology from an evolutionary perspective has proven an ideal occupational choice for a curious young person who couldn't decide what he wanted to be when he grew up. Indeed, reviewing the history of this perspective reignites my postadolescent awe at the number of fascinating questions about human social psychology yet to be asked.

References

- Ackerman, J.M., Shapiro, J.R., Neuberg, S.L., Kenrick, D.T., Becker, D.V., Griskevicius, V., Maner, J.K. and Schaller, M. (2006) They all look the same to me (unless they're angry): From out-group homogeneity to out-group heterogeneity . Psychological Science , 17, 836–840. <http://dx.doi.org/10.1111/j.1467-9280.2006.01790.x>

Ackerman, J.M., Becker, D.V., Mortensen, C.R., Sasaki, T., Neuberg, S.L. and Kenrick, D.T. (2009) A pox on the mind: Disjunction of attention and memory in the processing of physical disfigurement . *Journal of Experimental Social Psychology* 45, 478–485. <http://dx.doi.org/10.1016/j.jesp.2008.12.008>

Alcock, J. and Crawford, C. (2008) Evolutionary questions for evolutionary psychologists . In C. Crawford, ed. and D.L. Krebs (eds), *Foundations of Evolutionary Psychology* pp. 25–45. Mahwah NJ: Lawrence Erlbaum Associates.

Allen-Arave, W., Gurven, M. and Hill, K. (2008) Reciprocal altruism, rather than kin selection, maintains nepotistic food transfers on an Ache reservation . *Evolution and Human Behavior* , 29, 305–318. <http://dx.doi.org/10.1016/j.evolhumbehav.2008.03.002>

Archer, J. (2000) Sex differences in aggression between heterosexual partners: A meta-analytic review . *Psychological Bulletin* , 726, 651–680. <http://dx.doi.org/10.1037/0033-2909.126.5.651>

Becker, D.V., Kenrick, D.J., Guerin, S. and Maner, J.K. (2005) Concentrating on beauty: Sexual selection and sociospatial memory . *Personality and Social Psychology Bulletin* , 31, 1643–1652. <http://dx.doi.org/10.1177/0146167205279583>

Becker, D.V., Kenrick, D.J., Neuberg, S.L., Blackwell, K.C. and Smith, D.M. (2007) The confounded nature of angry men and happy women . *Journal of Personality and Social Psychology* , 92, 179–190. <http://dx.doi.org/10.1037/0022-3514.92.2.179>

Brewer, M.B. and Caporael, L.R. (2006) An evolutionary perspective on social identity: Revisiting groups . In M. Schaller, ed. , J. Simpson, ed. , and D.J. Kenrick, ed. , *Evolution and Social Psychology* , pp. 143–161. New York: Psychology Press.

Buss, D.M. (1989) Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures . *Behavioral and Brain Sciences* , 12, 1–49. <http://dx.doi.org/10.1017/S0140525X00023992>

Buss, D.M. (2000) The evolution of happiness . *American Psychologist* , 55, 15–23. <http://dx.doi.org/10.1037/0003-066X.55.1.15>

Cameron, C., Oskamp, S. and Sparks, W. (1977) Courtship American style – newspaper ads . Family Coordinator , 26, 27–30. <http://dx.doi.org/10.2307/581857>

Cialdini, R.B., Brown, S.L., Lewis, B.P., Luce, C. and Neuberg, S.L. (1997) Reinterpreting the empathy-altruism relationship: When one into one equals oneness . Journal of Personality and Social Psychology , 73, 481–494. <http://dx.doi.org/10.1037/0022-3514.73.3.481>

Cosmides, L. and Tooby, J. (1992) Cognitive adaptations for social exchange . In J.H. Barkow, ed. , L. Cosmides, ed. , and J. Tooby, ed. , The Adapted Mind , pp. 163–228. New York: Oxford University Press.

Cottrell, C.A. and Neuberg, S.L. (2005) Different emotional reactions to different groups: A sociofunctional threat-based approach to ‘prejudice’ . Journal of Personality and Social Psychology , 88, 770–789. <http://dx.doi.org/10.1037/0022-3514.88.5.770>

Crawford, C.B. and Krebs, D. (2008) Foundations of Evolutionary Psychology . New York: Erlbaum.

Crawford, C.B. and Salmon, C. (2004) Evolutionary Psychology, Public Policy, and Personal Decisions . Mahwah, NJ: Lawrence Erlbaum Associates.

Daly, M. and Wilson, M.I. (1983) Sex, Evolution and Behavior: Adaptations for Reproduction (2nd edition). Boston, MA: Willard Grant Press.

Daly, M. and Wilson, M.I. (1988) Homicide . New York: Aldine deGruyter.

Darwin, C. (1872) The Expression of Emotions in Man and Animals . London: Murray. <http://dx.doi.org/10.1037/10001-000>

Deci, E.L. and Ryan, R.M. (2000) The ‘what’ and ‘why’ of goal pursuits: Human needs and the self-determination of behavior . Psychological Inquiry , 11, 227–268. http://dx.doi.org/10.1207/S15327965PLI1104_01

Dunbar, R.I.M. and Barrett, L. (2007) The Oxford Handbook of Evolutionary Psychology . Oxford: Oxford University Press.

Eagly, A.H. and Wood, W. (1999) The origins of sex differences in human behavior: Evolved dispositions versus social roles . *American Psychologist* , 54, 408–423. <http://dx.doi.org/10.1037/0003-066X.54.6.408>

Eibl-Eibesfeldt, I. (1973) The expressive behavior of the deaf-and-blind-born . In M. von Cranach, ed. and I. Vine (eds), *Social Communication and Movement* , pp. 163–194. New York: Academic Press.

Eibl-Eibesfeldt, I. (1975) *Ethology: The Biology of Behavior* (2nd edition). New York: Holt, Rinehart, and Winston.

Ermer, E., Cosmides, L. and Tooby, J. (2008) Relative status regulates risky decision making about resources in men: Evidence for the co-evolution of motivation and cognition . *Evolution and Human Behavior* , 29, 106–118. <http://dx.doi.org/10.1016/j.evolhumbehav.2007.11.002>

Faulkner, J., Schaller, M., Park, J.H. and Duncan, L.A. (2004) Evolved disease-avoidance mechanisms and contemporary xenophobic attitudes . *Group Processes and Intergroup Relations* , 7, 333–353. <http://dx.doi.org/10.1177/1368430204046142>

Fiske, S.T., Cuddy, A.J.C. and Glick, P. (2007) Universal dimensions of social perception: Warmth and competence . *Trends in Cognitive Science* , 11, 77–83. <http://dx.doi.org/10.1016/j.tics.2006.11.005>

Gangestad, S.W., Garver-Apgar, C.E., Simpson, J.A. and Cousins, A.J. (2007) Changes in women's mate preferences across the ovulatory cycle . *Journal of Personality and Social Psychology* , 92, 151–163. <http://dx.doi.org/10.1037/0022-3514.92.1.151>

Gangestad, S.W. and Thornhill, R. (1998) Menstrual cycle variation in women's preferences for the scent of symmetrical men . *Proceedings of the Royal Society of London B* , 265, 727–733. <http://dx.doi.org/10.1098/rspb.1998.0380>

Garcia, J. and Koelling, R.A. (1966) Relation of cue to consequence in avoidance learning . *Psychonomic Science* , 4, 123–124.

Gardner, H. (1985) *The Mind's New Science: A History of the Cognitive Revolution*. New York: Basic Books.

Geary, D.C. (1998) *Male, Female: The Evolution of Human Sex Differences*. Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10370-000>

Griskevicius, V., Cialdini, R.B. and Kenrick, D.J. (2006a) Peacocks, Picasso, and parental investment: The effects of romantic motives on creativity. *Journal of Personality and Social Psychology*, 91, 63–76. <http://dx.doi.org/10.1037/0022-3514.91.1.63>

Griskevicius, V., Goldstein, N., Mortensen, C., Cialdini, R.B. and Kenrick, D.T. (2006b) Going along versus going alone: When fundamental motives facilitate strategic (non)conformity. *Journal of Personality and Social Psychology*, 91, 281–294. <http://dx.doi.org/10.1037/0022-3514.91.2.281>

Griskevicius, V., Tybur, J.M., Sundie, J.M., Cialdini, R.B., Miller, G.F. and Kenrick, D.J. (2007) Blatant benevolence and conspicuous consumption: When romantic motives elicit strategic costly signals. *Journal of Personality and Social Psychology*, 93, 85–102. <http://dx.doi.org/10.1037/0022-3514.93.1.85>

Harton, H.C. and Bourgeois, M.J. (2004) Cultural elements emerge from dynamic social impact. In M. Schaller, ed. and C. Crandall (eds), *The Psychological Foundations of Culture*, pp. 41–75. Mahwah, NJ: Lawrence Erlbaum Associates.

Haselton, M.G. and Buss, D.M. (2000) Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology* 78, 81–91. <http://dx.doi.org/10.1037/0022-3514.78.1.81>

Haselton, M.G. and Nettle, D. (2006) The paranoid optimist: An integrative evolutionary model of cognitive bias. *Personality and Social Psychology Review*, 10, 47–66. http://dx.doi.org/10.1207/s15327957pspr1001_3

Higgins, E.T. (1998) Promotion and prevention: regulatory focus as a motivational principle . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 30, 1–46. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960381-0>

Higgins, E.T. and Pittman, T.S. (2008) Motives of the human animal: Comprehending, managing, and sharing inner states . *Annual Review of Psychology* , 59, 361–385. <http://dx.doi.org/10.1146/annurev.psych.59.103006.093726>

Huang, J.Y. and Bargh, J.A. (2008) Peak of desire: activating the mating goal changes life-stage preferences across living kinds . *Psychological Science* . 19, 573–578. <http://dx.doi.org/10.1111/j.1467-9280.2008.02125.x>

Jankowiak, W.R. and Fischer, E.F. (1992) A cross-cultural perspective on romantic love . *Ethnology* , 31, 149–155. <http://dx.doi.org/10.2307/3773618>

Kaplan, H.S., Hill, K., Lancaster, J.B. and Hurtado, A.M. (2000) A theory of human life history evolution: Diet, intelligence, and longevity . *Evolutionary Anthropology* 9, 156–185. <http://dx.doi.org/10.1002/1520-6505%282000%299:4%3C156::AID-EVAN5%3E3.0.CO;2-7>

Keller, M.C. and Nesse, R.M. (2006) The evolutionary significance of depressive symptoms: Different adverse situations lead to different depressive symptom patterns . *Journal of Personality and Social Psychology* , 91, 316–330. <http://dx.doi.org/10.1037/0022-3514.91.2.316>

Kenrick, D.J. (1994) Evolutionary social psychology: From sexual selection to social cognition . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 26, 75–121. San Diego, CA: Academic Press.

Kenrick, D.J. (2001) Evolution, cognitive science, and dynamical systems: An emerging integrative paradigm . *Current Directions in Psychological Science* , 10, 13–17. <http://dx.doi.org/10.1111/1467-8721.00104>

Kenrick, D.J. (2006) Evolutionary psychology: Resistance is futile . *Psychological Inquiry* 17, 102–108.

Kenrick, D.J., Delton, A.W., Robertson, T., Becker, D.V. and Neuberg, S.L. (2007) How the mind warps: A social evolutionary perspective on cognitive processing disjunctions . In J.P. Forgas, ed. , M.G. Haselton, ed. , and W. Von Hippel (eds), *The Evolution of the Social Mind: Evolution and Social Cognition* , pp. 49–68. New York: Psychology Press.

Kenrick, D.J., Gabrielidis, C., Keefe, R.C. and Cornelius, J. (1996a) Adolescents' age preferences for dating partners: Support for an evolutionary model of life-history strategies . *Child Development* , 67, 1499–1511. <http://dx.doi.org/10.2307/1131714>

Kenrick, D.J., Griskevicius, V., Neuberg, S.L. and Schaller, M. (2010) Renovating the pyramid of needs: Contemporary extensions built upon ancient foundations . *Perspectives on Psychological Science* , 5, 292–314. <http://dx.doi.org/10.1177/1745691610369469>

Kenrick, D.J., Griskevicius, V., Sundie, J.M., Li, N.P., Li, Y.J. and Neuberg, S.L (2009) Deep rationality: The evolutionary economics of decision-making . *Social Cognition* , 27, 764–785. <http://dx.doi.org/10.1521/soco.2009.27.5.764>

Kenrick, D.J., Groth, G.R., Trost, M.R. and Sadalla, E.K. (1993) Integrating evolutionary and social exchange perspectives on relationships: Effects of gender, self-appraisal, and involvement level on mate selection criteria . *Journal of Personality and Social Psychology* , 64, 951–969. <http://dx.doi.org/10.1037/0022-3514.64.6.951>

Kenrick, D.J. and Keefe, R.C. (1992) Age preferences in mates reflect sex differences in human reproductive strategies . *Behavioral and Brain Sciences* , 15, 75–133. <http://dx.doi.org/10.1017/S0140525X00067595>

Kenrick, D.J., Li, N.P. and Butner, J. (2003) Dynamical evolutionary psychology: Individual decision-rules and emergent social norms . *Psychological Review* , 110, 3–28. <http://dx.doi.org/10.1037/0033-295X.110.1.3>

Kenrick, D.J., Neuberg, S.L., Zierk, K. and Krones, J. (1994) Evolution and social cognition: Contrast effects as a function of sex, dominance, and physical attractiveness . *Personality and Social Psychology Bulletin* , 20, 210–217. <http://dx.doi.org/10.1177/0146167294202008>

Kenrick, D.J., Nieuweboer, S. and Buunk, A.P. (2010) Universal mechanisms and cultural diversity: Replacing the blank slate with a coloring book . In M. Schaller, ed. , S. Heine, ed. , A. Norenzayan, ed. , T. Yamagishi, ed. and T. Kameda (eds), Evolution, Culture, and the Human Mind , pp. 257–271. New York: Psychology Press.

Kenrick, D.J., Sadalla, E.K., Groth, G. and Trost, M.R. (1990) Evolution, traits, and the stages of human courtship: Qualifying the parental investment model . Journal of Personality , 58, 97–116. <http://dx.doi.org/10.1111/j.1467-6494.1990.tb00909.x>

Kenrick, D.J. and Sheets, V. (1994) Homicidal fantasies . Ethology and Sociobiology , 14, 231–246. <http://dx.doi.org/10.1016/0162-3095%2893%2990019-E>

Kenrick, D.J. and Shiota, M.N. (2008) Approach and avoidance motivation(s): An evolutionary perspective . In A.J. Elliot (ed.), Handbook of Approach and Avoidance Motivation , pp. 273–288. New York: Psychology Press.

Kenrick, D.J., Sundie, J.M. and Kurzban, R. (2008) Cooperation and conflict between kith, kin, and strangers: Game theory by domains . In C. Crawford, ed. and D. Krebs (eds), Foundations of Evolutionary Psychology , pp. 353–370. New York: Lawrence Erlbaum Associates.

Kenrick, D.J., Trost, M.R. and Sheets, V.L. (1996b) The feminist advantages of an evolutionary perspective . In D.M. Buss, ed. and N. Malamuth (eds), Sex, Power, Conflict: Feminist and Evolutionary Perspectives , pp. 29–53. New York: Oxford University Press.

Kurzban, R., Tooby, J. and Cosmides, L. (2001) Can race be erased? Coalitional computation and social categorization . Proceedings of the National Academy of Sciences , 98, 15387–15392. <http://dx.doi.org/10.1073/pnas.251541498>

Lancaster, J.B. (1975) Primate Behavior and the Emergence of Human Culture . New York: Holt, Rinehart, and Winston.

Latané, B. (1996) Dynamic social impact: The creation of culture by communication . Journal of Communication , 46, 13–25.

Leary, M.R., and Baumeister, R.F. (2000) The nature and function of self-esteem: Sociometer theory . *Advances in Experimental Social Psychology* , 32, 1–62. <http://dx.doi.org/10.1016/S0065-2601%2800%2980003-9>

Li, N.P., Bailey, J.M., Kenrick, D.J. and Linsenmeier, J.A. (2002) The necessities and luxuries of mate preferences: Testing the trade-offs . *Journal of Personality and Social Psychology* , 82, 947–955. <http://dx.doi.org/10.1037/0022-3514.82.6.947>

Li, N.P. and Kenrick, D.J. (2006) Sex similarities and differences in preferences for short-term mates: What, whether, and why . *Journal of Personality and Social Psychology* , 90, 468–489. <http://dx.doi.org/10.1037/0022-3514.90.3.468>

Li, Y.J., Kenrick, D.J., Griskevicius, V. and Neuberg, S.L. (2010, June) The Evolutionary Roots of Decision Biases: Erasing and Exacerbating Loss Aversion . Paper presented at the conference for the Human Behavior and Evolution Society, Eugene, Oregon.

Lummaa, V. (2007) Life-history theory, reproduction and longevity in humans . In R.I.M. Dunbarand, ed. and L. Barrett (eds), *Oxford Handbook of Evolutionary Psychology* , pp. 397–413. Oxford: Oxford University Press.

Maner, J.K., Kenrick, D.J., Becker, D.V., Delton, A.W., Hofer, B., Wilbur, C.J. and Neuberg, S.L. (2003) Sexually selective cognition: Beauty captures the mind of the beholder . *Journal of Personality and Social Psychology* , 85, 1107–1120. <http://dx.doi.org/10.1037/0022-3514.85.6.1107>

Maner, J.K., Kenrick, D.J., Becker, D.V., Robertson, T.E., Hofer, B., Neuberg, S.L., Delton, A.W., Butner, J. and Schaller, M. (2005) Functional projection: How fundamental social motives can bias interpersonal perception . *Journal of Personality and Social Psychology* , 88, 63–78. <http://dx.doi.org/10.1037/0022-3514.88.1.63>

Mikulincer, M. and Shaver, P.R. (2006) The behavioral systems construct: A useful tool for building an integrative model of the social mind . In P.A.M. Van Lange (ed.), *Bridging Social Psychology* , pp. 279–284. Mahwah, NJ: Erlbaum.

Murstein, B.I. (1981) Process, filter, and stage theories of attraction . In M. Cook (ed.) *The Bases of Human Sexual Attraction* , pp. 179–211. San Diego: Academic Press.

Navarrete, C.D., Olsson, A., Ho, A.K., Mendes, W.B., Thomsen, L. and Sidanius, J. (2009) Fear extinction to an out-group face: The role of target gender . Psychological Science , 20, 155–158. <http://dx.doi.org/10.1111/j.1467-9280.2009.02273.x>

Neuberg, S.L., Kenrick, D.J. and Schaller, M. (2010) Evolutionary social psychology, pp. 761–796 in S.T. Fiske, ed. , D.J. Gilbert, ed. , and G. Lindzey (eds), Handbook of Social Psychology (5th edition). New York: John Wiley and Sons.

Nowak, A. and Vallacher, R.R. (1998) Dynamical Social Psychology New York: Guilford.

Ohman, A. and Mineka, S. (2001) Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning . Psychological Review , 108, 483–522. <http://dx.doi.org/10.1037/0033-295X.108.3.483>

Penn, D.J. (2003) The evolutionary roots of our environmental problems: Toward a Darwinian ecology . Quarterly Review of Biology , 78, 275–301. <http://dx.doi.org/10.1086/377051>

Penton-Voak, I.S., Little, A.C., Jones, B.C., Burt, D.M., Tiddeman, B.P. and Perrett, D.J. (2003) Female condition influences preferences for sexual dimorphism in faces of male humans (*Homo sapiens*) . Journal of Comparative Psychology , 117, 264–271. <http://dx.doi.org/10.1037/0735-7036.117.3.264>

Pillsworth, E.G. and Haselton, M.G. (2006) Male sexual attractiveness predicts differential ovulatory shifts in female extra-pair attraction and male mate retention . Evolution and Human Behavior , 27, 247–258. <http://dx.doi.org/10.1016/j.evolhumbehav.2005.10.002>

Pinker, S. (1994) The Language Instinct . New York: Harper Collins.

Roese, N., Pennington, G.L., Coleman, J., Janicki, M., Li, N.P. and Kenrick, D.T. (2006) Sex differences in regret: All for love or some for lust? Personality and Social Psychology Bulletin , 32, 770–780. <http://dx.doi.org/10.1177/0146167206286709>

Rowe, D.C. (1994) The Limits of Family Experience: Genes, Experience, and Behavior . New York: Guilford.

Sadalla, E.K., Kenrick, D.T. and Vershure, B. (1987) Dominance and heterosexual attraction . *Journal of Personality and Social Psychology* , 52, 730–738. <http://dx.doi.org/10.1037/0022-3514.52.4.730>

Schaller, M. and Murray, D.R. (2008) Pathogens, personality and culture: Disease prevalence predicts worldwide variability in sociosexuality, extraversion, and openness to experience . *Journal of Personality and Social Psychology* , 95, 212–221. <http://dx.doi.org/10.1037/0022-3514.95.1.212>

Schaller, M., Park, J.H. and Mueller, A. (2003) Fear of the dark: Interactive effects of beliefs about danger and ambient darkness on ethnic stereotypes . *Personality and Social Psychology Bulletin* , 29, 637–649. <http://dx.doi.org/10.1177/0146167203029005008>

Schaller, M., ed. , Simpson, J., ed. and Kenrick, D.T. (eds) (2006) Evolution and Social Psychology . New York: Psychology Press.

Schmitt, D.P. and 118 members of International Sexuality Description Project. (2003) Universal sex differences in the desire for sexual variety: Tests from 52 nations, 6 continents, and 13 islands . *Journal of Personality and Social Psychology* , 85, 85–104. <http://dx.doi.org/10.1037/0022-3514.85.1.85>

Sedikides, C. and Skowronski, J.J. (1997) The symbolic self in evolutionary context . *Personality and Social Psychology Review* , 7, 80–102. http://dx.doi.org/10.1207/s15327957pspr0101_6

Segerstale, U. (2000) Defenders of the Truth: The Battle for Science in the Sociology Debate and Beyond . Oxford: Oxford University Press.

Seligman, M.E.P. and Hager, J.L. (1972) Biological Boundaries of Learning . New York: Appleton-Century-Crofts.

Sherman, P.W. (1988) The levels of analysis . *Animal Behavior* , 36, 616–619. <http://dx.doi.org/10.1016/S0003-3472%2888%2980039-3>

Sherry, D.F. and Schacter, D.L. (1987) The evolution of multiple memory systems . Psychological Review , 94, 439–454. <http://dx.doi.org/10.1037/0033-295X.94.4.439>

Sidanius, J. and Pratto, F. (1999) Social Dominance: An Intergroup Theory of Social Hierarchy and Oppression . New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139175043>

Skinner, B.F. (1953) Science and Human Behavior . New York: Free Press.

Stearns, S.C., Allai, N. and Mace, R. (2008) Life history theory and human development . In C. Crawford, ed. and D. Krebs (eds), Foundations of Evolutionary Psychology , pp. 47–70. New York: Lawrence Erlbaum Associates.

Taylor, S.E., Klein, L.C., Lewis, B.P., Gruenwald, T.L., Gurung, R.A.R. and Updegraff, J.A. (2000) Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight . Psychological Review , 107, 411–429. <http://dx.doi.org/10.1037/0033-295X.107.3.411>

Trivers, R.L. (1972) Parental investment and sexual selection . In B. Campbell (ed.) Sexual Selection and the Descent of Man , pp. 136–179. Chicago: Aldine-Atherton.

Tversky, A. and Kahneman, D. (1991) Loss aversion in riskless choice: A reference dependent model . Quarterly Journal of Economics , 106, 1039–1061. <http://dx.doi.org/10.2307/2937956>

Tybur, J.M., Miller, G.F. and Gangestad, S.W. (2008) Testing the controversy: An empirical examination of adaptationists' attitudes toward politics and science . Human Nature: An Interdisciplinary Biosocial Perspective , 18, 313–328. <http://dx.doi.org/10.1007/s12110-007-9024-y>

Van Knippenberg, A. and Van Baaren, R. (2006) Baboons, brains, babies, and bonding: A multidisciplinary approach to mimicry . In P.A.M. Van Lange (ed.), Bridging Social Psychology , pp. 173–178. Mahwah, NJ: Erlbaum.

Van Vugt, M. and Van Lange, P.A.M. (2006) The altruism puzzle: Psychological adaptations for prosocial behavior . In M. Schaller, ed. , J.A. Simpson, ed. , and

D.J. Kenrick (eds), *Evolution and Social Psychology* , pp. 237–261. Madison, CT: Psychology Press.

Weeden, J., Cohen, A.B. and Kenrick, D.T. (2008) Religious attendance as reproductive support . *Evolution and Human Behavior* , 29, 327–334. <http://dx.doi.org/10.1016/j.evolhumbehav.2008.03.004>

Wilcoxon, H.C., Dragoin, W.B. and Krai, P.A. (1971) Illness-induced aversions in rat and quail: Relative salience of visual and gustatory cues . *Science* , 171, 826–828. <http://dx.doi.org/10.1126/science.171.3973.826>

Wilson, M. and Daly, M. (2003) Do pretty women inspire men to discount the future? *Proceedings of the Royal Society London B* , 271, S177–S179.

Wilson, E.O. (1975) *Sociobiology* . Cambridge, MA: Harvard University Press.

Wilson, E.O. (1998) *Consilience: The Unity of Knowledge* . New York: Alfred A. Knopf.

<http://dx.doi.org/10.4135/9781446249215.n2>

Handbook of Theories of Social Psychology: Volume 1

Tend and Befriend Theory

Contributors: Shelley E. Taylor

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Tend and Befriend Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n3>

Print pages: 32-50

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n3>

[p. 32 ↓]

Chapter 2: Tend and Befriend Theory

Shelley E. Taylor, ed.

Abstract

The “tend and befriend” theory builds on the observation that human beings affiliate in response to stress. Under conditions of threat, they tend to offspring to ensure their survival and affiliate with others for joint protection and comfort. These responses are underpinned by an affiliative neuro-circuitry that appears to be based on oxytocin and endogenous opioid peptides. When close relationships are threatened or one is socially isolated, a rise in plasma oxytocin occurs, a biological marker that may signal a need for affiliation. Oxytocin prompts affiliative behavior in response to stress, in conjunction with the opioid system. Together with positive social contacts, oxytocin attenuates biological stress responses that would otherwise arise in response to social threats. These social responses to stress and their biological underpinnings appear to be more characteristic of women than men. This model helps to unravel puzzles not only in the research literature but also with respect to health and may shed light on why women tend to live longer than men.

Introduction

Social relationships are vital resources for managing the demands of the environment. Human beings' social nature may have evolved in part because of the needs that we have for contact with others to ensure safety from threat. Whereas other species have thick skin, sharp teeth, quick reflexes, or camouflage to protect themselves, human

beings have adopted group living as the primary solution to problems of survival and reproduction. Social isolation is both psychologically and physically toxic, and it is associated with a heightened risk of death and early mortality (Cacioppo and Hawkley, 2003; House et al., 1988). In addition to survival benefits, social relationships serve important regulatory functions. For example, contact with a nurturant caregiver early in life is essential for the development of biological stress regulatory systems (e.g., Repetti et al., 2002, 2007), and social support is protective across the lifespan (see Taylor, 2009, for a review).

Responses to Threat: Fight or Flight

Despite the centrality of the social group to human wellbeing and survival, research on stress and threat has, until recently, largely [p. 33 ↓] ignored the importance of social mechanisms for addressing responses to threat. Instead, stress research has been guided heavily by the fight or flight metaphor, responses that may protect individuals but that are not social in nature. First articulated by Walter Cannon (1932), the fight or flight response has two components: a behavioral component and a biological component. In response to a threat, a person can become aggressive and mount an active or antagonistic response to the threatening circumstances, or the person can flee, either literally or metaphorically. Among the responses that contemporary stress researchers interpret as flight behavior are social withdrawal and substance use, especially alcohol and drug use.

The biological component of fight or flight depends on two interacting stress systems: the sympathetic nervous system (SNS) and the hypothalamic pituitary adrenocortical (HPA) axis. The actions of the sympathetic adrenomedullary (SAM) system are mediated primarily by the catecholamines norepinephrine and epinephrine, which produce, among other changes, increased heart rate and blood pressure, dilation of the airways, and enhanced availability of glucose and fatty acids for energy. Engagement of the HPA axis begins with the release of corticotropin-releasing hormone (CRH), which stimulates the secretion of adrenocorticotropin hormone (ACTH) by the anterior pituitary gland, resulting in the release of glucocorticoids such as Cortisol. Cortisol helps to restore processes that prime homeostatic defense mechanisms, and as such, HPA axis

reactivity modulates a wide range of functions, including energy release and immune activity.

These two systems are important because they account for the protective short-term effects of stress responses but also their long-term costs. In the short term, they shunt reserves of energy for fight or flight, and the subjective experience is often arousal and feelings of anger, fear, or anxiety. As such, these systems mobilize the body to meet the demands of pressing situations and then engage homeostatic mechanisms that return the body to its previous functioning. With repeated or recurrent stress, however, these biological stress responses have long-term costs that have implications for health (McEwen, 1998). The theory of allostatic load (McEwen, 1998) maintains that repeated or chronic engagement of stress systems interacts with genetic vulnerabilities and acquired risks (such as poor health habits) to erode the resiliency of biological stress regulatory systems and increase the likelihood of disease. These include such chronic disorders as coronary heart disease, hypertension, type II diabetes, and some cancers. These long-term costs provide important conceptual underpinnings to the tend and befriend theory to be articulated, because, as will be seen, social contact during times of stress and social relationships more generally exert protective effects against these potential long-term costs of stress.

When stress researchers began to study stress in human beings, they borrowed from animal research in ways conducive to identifying fight or flight responses in humans (Taylor et al., 2000). Although fighting and fleeing are unquestionably part of the human repertoire for responding to threat, there are at least two reasons to suspect that they are unlikely to be the only or primary responses. First, fighting can leave vulnerable offspring at risk for predation. Likewise, fleeing with an infant or toddler might slow the caregiver down to be at enhanced risk for attack.

Human beings would not have survived as a species had they not developed stress responses that protected their offspring as well as self in times of danger. Humans evolved in small hunter-gatherer groups, and so coming together as a group instead of fighting or fleeing would provide greater defense as well as information about resource location for combating threats. In short, there are many reasons to believe that humans have used social relationships not only as a basic accommodation to the exigencies of life, but also as a primary resource for dealing with stressful circumstances.

[p. 34 ↓]

Responses to Threat: Tend and Befriend

To characterize these social responses to threat, we developed the theory “tend and befriend” (Taylor, 2002; Taylor et al., 2000). The theory maintains that under conditions of threat, tending to offspring and affiliating with others (befriending) are common responses in humans.

Development of the Theory

The theory of tend and befriend was initiated in 1998 when the members of our laboratory group went to hear a prominent animal stress researcher present his work. He made the statement: “We shocked the animals and, of course, they all attacked each other.” This was an arresting statement, in large part because it is quite contrary to what one sees in human beings. When there is an immediate threat, human beings are more likely to affiliate with one another and to offer each other aid or solace than to attack one another. Accordingly, our lab group began a series of discussions to develop what would become the theory of tend and befriend as a characterization of human social responses to stress.

Our point of departure was evolutionary theory. A dilemma with any evolutionary-based theory of social behavior, however, is the potential for “just so” stories, a reference to Rudyard Kipling’s fables about how the leopard got his spots, the elephant his trunk, and the like. Such stories can seem plausible, but have no basis in fact. Accordingly, in building our theory, we imposed constraints on its development. Our method of constraint was to build parallel and mutually constraining biological and behavioral models. For every hypothesis we developed that generated a behavioral statement for the tend and befriend theory, we required evidence at the biological level. The reverse was also true. Any biological literature that yielded a potential insight for the theory was constrained by a requirement of evidence at the behavioral level. Using these criteria, we jettisoned a number of otherwise promising hypotheses early on. Thus, the following

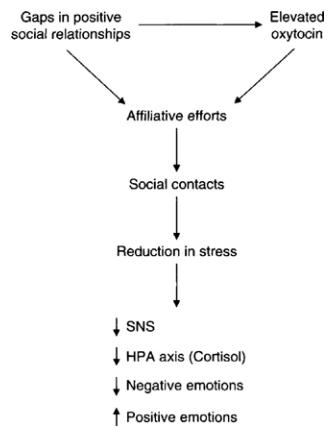
characterization of the theory meets the criteria for evidence at both the biological and behavioral levels.

Overview of the Theory

Drawing on animal and human studies, we maintain that there is an affiliative neurocircuitry that monitors the adequacy of social contact in light of the demands of the environment and prompts affiliation when necessary, regulating social approach behavior. We suggest that this system works in much the same way as occurs for other appetitive needs. That is, just as people have basic needs of hunger, thirst, and sexual drive, they also need to maintain an adequate level of protective and rewarding social relationships.

We hypothesize that there is a biological signaling system that comes into play when affiliations fall below an adequate level. Once signaled, the appetitive need is met through purposeful social behavior, such as affiliation. If social contacts are hostile and unsupportive, then psychological and biological stress responses are heightened and efforts toward social affiliation may be redoubled. If social contacts are supportive and comforting, stress responses decline. These positive contacts, then, lead to a decline in the need for affiliation and, in the context of stress, a decline in biological stress responses. The model is pictured in [Figure 2.1](#).

Figure 2.1 Tend and befriend



On the biological level, the theory draws heavily on oxytocin and the opioid system. Oxytocin and endogenous opioid peptides are released in response to at least some stressors, especially those that trigger affiliative needs. Oxytocin prompts affiliative behavior in response to stress in conjunction with the opioid system, and oxytocin together with positive social contacts attenuates biological stress responses that otherwise arise in response to social threats. We suggest that [p. 35 ↓] this oxytocin-opioid system is an appetitive system that regulates social approach behavior and recruits the neurocircuitry for reward in its enactment. Finally, we maintain that some of the health benefits associated with social support and social integration are mediated by this appetitive social approach system via attenuation of threat responses.

Tend and befriend responses to stress may be particularly characteristic of women. At the time when human stress responses evolved, generally thought to be the Pleistocene Era, tasks of daily living were largely sex-segregated, with men heavily responsible for protection and hunting and women primarily responsible for childcare and foraging. Consequently, women's responses to threat would have evolved so as to protect not only self but also immature offspring in their care. Affiliating with the social group for joint protection of self and offspring would have had substantial survival benefits, helping to ensure that offspring would reach an age when they could reproduce.

This is not to suggest that tend and befriend responses are exclusive to women. A large literature indicates that under stress, both men and women turn to others for protection and solace (see Taylor, 2009, for a review). A gender difference exists such that women

are somewhat more likely to seek and use social support in response to stress than men are (Tamres et al., 2002). However, this difference, although robust, is relatively modest in size, and thus, men's social responses to stress are well documented.

Theoretical Principles and Evidence for Tend and Befriend

Biological Signaling System: Social Pain/Separation

Because affiliation is vital to the survival of human beings, there are likely to be biobehavioral mechanisms that are sensitive to social threats or loss of social contact, resulting in social distress and consequent efforts to establish or restore positive social contacts. One paradigm for such a system is separation distress, which has been studied primarily in young animals and human infants. When the young are separated from the caregiver, separation distress can result leading to distress vocalizations (e.g., crying in human infants) or active efforts to find the caregiver.

This system appears to depend in part on brain opioids. Evidence consistent with this pathway includes the fact that brain opioids reduce separation distress and opioid-based drugs, such as morphine, reduce distress vocalizations in animal offspring separated from the mother (Panksepp, 1998). Similarly, depriving animals of companionship can increase their consumption of exogenous opioids. Genetic evidence, likewise, suggests a role for opioids in the separation distress process. Mice that lack the μ -opioid receptor gene emit fewer distress vocalizations when separated from their mothers (Moles et al., 2004). Of interest, the opioid system is also vital to the experience and management of physical pain, and so researchers have [p. 36 ↓] inferred that the experience of social pain and separation distress may build on the neurocircuitry for physical pain. Recent genetic and neuroimaging studies with humans have lent credibility to this idea (Eisenberger et al., 2003; Way and Taylor, 2011).

Oxytocin also appears to be implicated in distress due to social isolation. Adults as well as young children encounter gaps in their social relationships and may experience an analogue of separation distress. To address this point, in a laboratory study, we gave women measures of psychological and social functioning and related their responses to their levels of plasma oxytocin (Taylor et al., 2006a). The questionnaires assessed gaps in the women's relationships, including recent declines in contacts with significant others and how positive and negative their relationships were. Women who were experiencing gaps in their social relationships had high levels of oxytocin. They were more likely to report reduced contact with their mothers, with their best friend, with a pet, and with the social groups to which they belonged. Oxytocin levels were also related to the absence of positive relations with a partner. Specifically, women who reported that their husbands were not supportive, did not understand how they felt about things, and did not care for them, had higher levels of oxytocin. These women reported that they could not open up to their husbands if they needed to share their concerns. Poor quality of the marital relationship and infrequent displays of affection were also associated with high levels of oxytocin. These findings suggest that oxytocin is sensitive to the absence of positive significant social relationships. Similar results were reported by Turner et al. (1999) who found that elevated oxytocin was associated with anxiety over relationships, not being in a primary relationship, or having cold, intrusive relationships (see also Taylor et al., 2009).

It is of course possible that women with high levels of oxytocin are inclined to construe their social relationships as unsupportive. Some evidence argues against this direction of causality: women who reported declines in contact with a pet and mother often had experienced their deaths, and women's oxytocin levels are unlikely to have caused these deaths. A recent animal study provides more definitive evidence regarding the direction of causality. Specifically, Grippo et al (2007) isolated female prairie voles and found that oxytocin levels increased in response to social isolation. In summary, social pain and separation are psychologically distressing, and the opioid and oxytocin systems are implicated in these responses.

Affiliation

The tend and befriend theory maintains that in response to either a psychological or biological impetus to affiliate or both, people seek contact with others. As an affiliative hormone, oxytocin may provide this impetus for social contact. Manifold evidence from animal studies shows that exogenous administration of oxytocin leads to affiliation in species as varied as rats, monkeys, and sheep. The injection of oxytocin leads to increases in maternal behavior, in grooming, and in other prosocial behavior (e.g., see Carter et al., 1999, for a review).

Opioid mechanisms also appear to be implicated in these processes. Administration of an opioid antagonist, for example, naloxone or naltrexone, results in less caregiving and protective behavior toward infants in rhesus monkeys (Martel et al., 1993), inhibits maternal behavior in sheep (Kendrick and Keverne, 1989), and diminishes the rewarding aspect of maternal cues in rats (Panksepp et al., 1999). Animals demonstrate a preference for other animals in whose presence they have previously experienced high levels of oxytocin and opioid activity (Panksepp, 1998). Administration of an opioid antagonist can suppress juvenile social behavior (Jalowiec et al., 1989), and opioid-blocking agents have been associated with reduced social activity and grooming in rhesus monkeys (Martel et al., 1993).

[p. 37 ↓] There is suggestive evidence that opioid-blocking agents may suppress human social behavior as well. Specifically, in one study (Jamner et al., 1998), administration of an opioid-blocking agent increased the amount of time that women chose to spend alone, reduced the amount of contact they had with friends, reduced the likelihood that they would contact their friends, and reduced the pleasantness of interactions with friends. Exogenous administration of oxytocin appears to enhance prosocial behavior and instill a sense of trust (Kosfeld et al., 2005; Zak et al., 2004). Thus, a broad array of affiliative behaviors appear to be subserved by oxytocin and opioid mechanisms in both animals and humans, although the animal data are more plentiful at this point in time.

Tending and Responses to Stress

The original tend and befriend model maintained that maternal nurturance and tending under stressful conditions have specific advantages for offspring by protecting them from physical harm and increasing the likelihood that they will grow to adulthood and reproduce. Recent empirical developments indicate that the importance of tending to offspring in response to threat is far broader than the early theoretical statements would imply.

Animal Studies

In an important animal model, Meaney and colleagues (Francis et al., 1999; Liu et al., 1997) explicitly linked early nurturant maternal contact following a stressful encounter to the development of stress responses in offspring. In their paradigm, infant rats are removed from the nest, stroked, and then returned to the nest. The response of the mother to this separation and reunification is intense licking and grooming and arched-back nursing, which provides the pup with nurturant, soothing immediate stimulation. On the short term, this contact reduces SAM and HPA axis responses to stress in the pups.

Over the long term, this maternal behavior results in a better regulated HPA axis response to stress and novelty, and better regulation of somatic growth and neural development, especially hippocampal synaptic development in the pup. These rat pups also show more open field exploration, which suggests lower levels of fear. This compelling animal model suggests that nurturant stimulation by the mother in response to stressful encounters modulates the responses of offspring to stress in ways that have permanent effects on the offspring's HPA axis responses to stress, on behavior suggestive of anxiety/fearfulness, and on cognitive function (see also Suomi, 1999).

Human Studies

Warm, nurturant, and supportive contact with a caregiver affects physiological and neuroendocrine stress responses in human infants and children just as in these animal

studies. Early research on orphans reported high levels of emotional disturbance, especially depression, in infants who failed to receive nurturant stimulating contact from a caregiver (Spitz and Wolff, 1946). More recent findings from Eastern European abandoned infants confirm that, without the affectionate attentions of caregivers, infants may fail to thrive and many die (Carlson and Earls, 1997).

Building on observations such as these, attachment theory characterizes how vital early nurturant contact is to psychological and biological development (Bowlby, 1969/1982). Through the comforting behavior of attachment figures, typically parents, infants learn to understand and respond to the world. Bowlby regarded attachment behavior as regulated by an innate motivational system to ensure physical and psychological proximity to a caregiver. When a child experiences the attachment figure as responsive, effective self-regulation, exploratory behavior, and [p. 38 ↓] normative biological responses to challenges may result. However, when there is a threat to this relationship, the child may react biologically and behaviorally with signs of stress and seek attention and comfort.

Attachment also moderates biological responses to stress. Studying 15-month-old children receiving well-baby examinations, Gunnar and her associates found that securely attached infants were less likely to show elevated Cortisol responses to normal stressors such as inoculations than less securely attached infants (Gunnar et al., 1996; see also Nachmias et al., 1996). The protective effects of secure attachment were especially evident for socially fearful or inhibited children (see also Hart et al., 1996; Levine and Wiener, 1988; see Collins and Feeney, 2000, for a discussion of attachment in adult supportive relationships).

Early on, children depend critically on physical contact with primary caregivers, but over time they develop internal working models of these interactions that may buffer them when the primary caregiver is absent. If significant others are perceived to be warm, responsive, and available, a secure attachment will result. If significant others are cold, rejecting, unpredictable, or insensitive, however, an anxious or insecure attachment may result. Instead of being able to draw on an internal working model of caregivers for comfort and solace, the child may instead regulate his or her behavior by withdrawing from others or excessively demanding attention.

To a degree, the attachments laid down early in life provide a model for adult attachments (Fraley, 2002). That is, warm, nurturant contact with parents provides a model not only for stress responses and emotion regulation throughout childhood, but also for adult social relationships, and children who come from families in which they experienced warm, attentive behavior are more likely to develop the social skills that serve them well across their lifespan (Repetti et al., 2002). A broad array of evidence demonstrates that children from supportive families are more likely than those from unsupportive families to form secure attachments, and to develop effective emotion regulation skills and social competencies that help them to regulate their responses to stress (Repetti et al., 2002).

In essence, then, the early family environment may provide the groundwork for emotion regulation skills and social competencies for managing stress across the lifespan. In families that are warm and nurturant, children develop secure attachments and learn to manage threat effectively with a lesser physiological/neuroendocrine toll. If they are raised in non-nurturant or conflict-ridden families, children instead experience threatening events more commonly and learn fewer socioemotional skills for managing stress.

Biological Mechanisms

Families characterized by unsupportive relationships have damaging outcomes for the mental, physical, and social health of their offspring, not only in the short term, but across the lifespan. Overt family conflict, manifested in recurrent episodes of anger and aggression, deficient nurturing, and family relationships that are cold, unsupportive, and/or neglectful have been associated with a broad array of adverse mental and physical health outcomes long into adulthood (Repetti et al., 2002, 2007). The chronic stress of unsupportive families produces repeated or chronic SNS activation in children, which in turn may lead to wear and tear on the cardiovascular system.

Epigenetic factors appear to be involved in these pathways as well. That is, maternal nurturance can induce long-lasting changes in the function of genes, which is an additional mechanism by which experiences of early nurturance can induce long-term behavioral alterations in emotional and social functioning. For example, Suomi (1987)

reported that highly reactive monkeys cross-fostered to nurturant mothers develop good socioemotional skills and achieve high status in the [p. 39 ↓] dominance hierarchy, whereas monkeys with reactive temperaments who are peer-raised develop poor socioemotional skills and end up at the bottom of the dominance hierarchy.

Such long-term effects of maternal care appear to be a result of epigenetic structural alterations (methylation) to the glucocorticoid receptor gene that affect its expression throughout the lifespan (Meaney and Szyf, 2005). Mothers showing high levels of nurturant behavior exhibit greater increases in oxytocin receptors during pregnancy, which is thought to trigger maternal responsivity (Meaney, 2001), and have higher levels of dopamine release when caring for their pups (Champagne et al., 2004). This especially nurturant mothering triggers greater increases in serotonin turnover in the pup, which initiates the cascade leading to the altered glucocorticoid receptor expression that affects adult reactivity to stress (Meaney and Szyf, 2005).

Related evidence has been uncovered with humans. For example, the harshness or nurturance of the early family environment is implicated in the expression of the serotonin transporter gene (5-HTTLPR). People with two copies of the short allele (short/short) of this gene and who have experienced childhood maltreatment are more likely to be diagnosed with major depressive disorder than individuals with one or two copies of the long allele who have experienced similar environments (Caspi et al., 2003; Kaufman et al., 2004). However, a study from our laboratory (Taylor et al., 2006b) qualifies this conclusion: the short allele may function not only as a risk allele for depression in conjunction with an adverse early environment, but as an allele reflecting general sensitivity to the environment, providing protection from symptoms of depression when the environment is nurturant. Using a nonclinical sample of 118 adult men and women, we assessed nurturance of the early family environment, depressive symptomatology, and 5-HTTLPR genotype. As expected, a stressful early family environment by itself was significantly related to depressive symptomatology. However, a significant gene-by-environment interaction between the 5-HTTLPR and the nurturance of the early family environment qualified this risk. Specifically, individuals with two copies of the short allele had greater depressive symptomatology if they had experienced early familial adversity compared to participants with the short/long or long/long genotypes, but significantly less depressive symptomatology if they reported a nurturant early environment.

Thus, long-term, often permanent effects of early nurturance are evident not only at the behavioral level, but also at the biological level and can include the functioning of relevant genes. Tending to offspring in times of stress, then, offers not only immediate protection, but also long-term protection in the form of biological and behavioral responses to stress.

Befriending and Responses to Stress

Just as offspring are benefited through tending, so befriending confers stress regulatory benefits.

Animal Studies

Animal studies with rats, sheep, prairie voles, and other species show that exogenous administration of oxytocin or stimulation of oxytocin secretion via stroking decreases sympathetic reactivity, blood pressure, pain sensitivity, and glucocorticoid levels, among other findings suggestive of reduced biological stress responses (Carter, 1998; Insel, 1997; Petersson et al., 1996; Uvnäs-Moberg, 1997; Uvnäs-Moberg et al., 1994). Oxytocin also reduces psychological distress, having anxiolytic properties (McCarthy, 1995). For example, exogenous administration of oxytocin enhances sedation and relaxation, is tied to signs of reduced fearfulness in rodent studies, and is tied to enhanced exploratory [p. 40 ↓] behavior (Uvnäs-Moberg et al., 1994; Mantella et al., 2003; McCarthy, 1995).

Human Studies

Similar findings have been identified in human studies. For example, high levels of oxytocin or exogenous administration of oxytocin in humans produce decreases in sympathetic activity (e.g., Light et al., 2000; Uvnäs-Moberg, 1997) and inhibit secretion of ACTH and Cortisol (Altemus et al., 1995; Chiodera and Legros, 1981). Heinrichs and colleagues (Heinrichs et al., 2003) found that exogenous administration of oxytocin produced lower anxiety and lower Cortisol levels during a laboratory stress

challenge; the reduced Cortisol response was especially pronounced in men who also experienced social support from a friend (see also Kosfeld et al., 2005; Zak et al., 2004). Breastfeeding mothers in whom oxytocin levels are high have lower anxiety, depression, and stress following breastfeeding as compared with bottle-feeding (Modahl and Newton, 1979). Oxytocin increases the sensitivity of brain opioid systems and so, when oxytocin injection is accompanied by an opioid-blocking agent, Cortisol levels do not change. Thus, some of the anti-stress pathways properties of oxytocin are probably mediated via an opioid pathway.

Social Support

When people affiliate in response to stress, they commonly experience social support. Social support is defined as the perception or experience that one is loved and cared for by others, esteemed and valued, and part of a social network of mutual assistance and obligations (Wills, 1991). Social support may come from a partner, relatives, friends, coworkers, social and community ties, strangers, and even a devoted pet (Allen et al., 2002).

Taxonomies of social support typically classify it into several specific forms. Informational support occurs when one person helps another to understand a stressful event better by providing information about the event. Instrumental support involves the provision of tangible assistance, such as services, financial assistance, and other specific aid or goods. Emotional support involves providing warmth and assistance to another person and reassuring that person that he or she is a valuable person for whom others care. Social support may involve the reality of using the social network for benefits such as these, but it can also involve simply the perception that such resources are available should they be needed. That is, just knowing that one is cared for and that one could obtain support from others is often comforting in its own right.

Social contacts and social support are psychologically beneficial. Social support reduces psychological distress such as depression or anxiety during times of stress (e.g., Fleming et al., 1982; Sarason et al., 1997; Lin et al., 1999). It promotes psychological adjustment to chronically stressful conditions, such as coronary artery disease (Holahan et al., 1997), diabetes, HIV (Turner-Cobb et al., 2002), and cancer

(Penninx et al., 1998; Stone et al., 1999), among many other health-related disorders. Social support protects against cognitive decline in older adults (Seeman et al., 2001), heart disease among the recently widowed (Sorkin et al., 2002), and psychological distress in response to traumatic events, such as 9/11 (Simeon et al., 2005), among other psychological benefits.

Health Benefits of Social Support

Social support has been tied to a variety of specific health benefits among individuals sustaining health risks. These include fewer complications during pregnancy and childbirth (Collins et al., 1993), less susceptibility to herpes attacks among infected individuals (VanderPlate et al., 1988), lower rates of myocardial infarction among individuals with diagnosed disease, a reduced likelihood of mortality from myocardial infarction [p. 41 ↓] (Kulik and Mahler, 1993; Wiklund et al., 1988), faster recovery from coronary artery disease surgery (King et al., 1993; Kulik and Mahler, 1993), better diabetes control (Marteau et al., 1987), longer survival in patients with end-stage renal disease (Cohen et al., 2007), and less pain among arthritis patients (Brown et al., 2003b).

Social support also contributes to survival. In a classic study of 7,000 community residents in Alameda County, CA, epidemiologists Lisa Berkman and Leonard Syme (1979) found that people who lacked social and community ties over the previous nine years were more likely to die of all causes during the follow-up period than those who cultivated or maintained their social relationships. Having social contacts predicted an average 2.8 years increased longevity among women and 2.3 years among men, and these differences persisted after controlling for socioeconomic status, health status at the beginning of the study, and health habits (Berkman and Syme, 1979; see also Rutledge et al., 2004).

The positive impact of social contacts on health is as powerful or more powerful a predictor of health and longevity than well-established risk factors for chronic disease and mortality, with effect sizes on par with smoking, blood pressure, lipids, obesity, and physical activity (House et al., 1988). And as noted, in both animal and human studies,

social isolation is tied to a significantly enhanced risk of mortality (House et al., 1988) and a heightened risk of both chronic and acute health disorders (Taylor, 2009).

Mechanisms Underlying Health Benefits

Although not all the mechanisms explaining these strong relationships are known, one key pathway is via stress responses (Cacioppo and Hawkley, 2003). When humans are socially isolated, their sympathetic nervous system and HPA axis responses to stress may continue unabated. Consistent with the theory of allostatic load described earlier, to the extent that contact with others in times of stress reduces sympathetic and HPA axis activity in response to threats, cumulative wear and tear on biological functioning is lessened. Social contact, then, may leave people less vulnerable to immunologic compromise in response to stress and to health disorders tied to the excessive or recurrent functioning of the sympathetic nervous system and HPA axis.

Whether the attenuation of stress responses by oxytocin and opioids contribute to these clinical effects of social support is, at present, unclear. However, animal research using a wound-healing paradigm suggests that this is a promising avenue for research (Detillion et al., 2004). In this study, female Siberian hamsters received a cutaneous wound and were then exposed to immobilization stress. The stressor increased Cortisol concentrations and impaired wound healing, but only in socially isolated and not in socially housed animals. Thus, social housing acted as a stress buffer. Removing Cortisol via adrenalectomy eliminated the impact of the stressor on wound healing, thereby implicating the HPA axis in the wound healing process. Of particular relevance to the current arguments, treating the isolated hamsters with oxytocin eliminated the stress-induced increases in Cortisol and facilitated wound healing; treating socially housed hamsters with an oxytocin antagonist delayed wound healing. These data strongly imply that social contacts protect against the adverse effects of stress through a mechanism that implicates oxytocin-induced suppression of the HPA axis. Thus, there appear to be discernible clinical consequences (wound healing) of oxytocin in conjunction with social contact.

To summarize, evidence that social responses to threat in the form of tending and befriending are associated with beneficial mental and physical health outcomes is

overwhelming. Both animal and human studies attest not only to the beneficial effects of social contact in times of stress, but also to the mechanisms that may underpin these relations.

[p. 42 ↓]

Sex Differences in Tend and Befriend

Most, but not all, of the research demonstrating both the psychological effects of oxytocin and endogenous opioid peptides as well as their effects on downregulating stress responses has been conducted with female animals or with women. Although there is evidence that these processes may be implicated in men's reduced stress responses as well (e.g., Heinrichs et al., 2003), the research is less plentiful. Moreover, the biological underpinnings of the theory would appear to be more consistent with what is known about women's hormonal profiles than men's. For example, oxytocin's effects are enhanced in the presence of estrogen (see Taylor et al., 2000).

There is, however, a hormone, vasopressin, that is very similar in molecular structure to oxytocin and whose effects appear to be enhanced in the presence of androgens, and so it may play a parallel role in male social behavior (Panksepp, 1998). Vasopressin is important to stress responses because it is involved in the maintenance of plasma volume and blood pressure during shock, among other functions. In certain monogamous species, most notably the prairie vole, it has also been tied to males' prosocial responses to stress, for example, guarding and patrolling of territory, defense of mate, and defense of offspring against intruders. The vasopressin receptor gene has also been tied to pair bonding, monogamous behavior (Lim et al., 2004), empathy, and altruistic behavior (Anckarsäter and Cloninger, 2007; Knafo et al., 2008). In a recent test of the potential role of vasopressin in men's social behavior, Taylor et al. (2009) examined whether elevations in vasopressin and oxytocin were associated with dissatisfaction in the pair-bond relationship in men and women. Consistent with previous research described earlier, oxytocin was elevated in women experiencing distress, but vasopressin was not. Exactly the reverse pattern was found for men, such that men who were in distressing pair bond relationships had elevated vasopressin but not oxytocin. Thus, elevated plasma vasopressin in men may act as a signal that

the pair bond relationship is jeopardized, just as elevations in plasma oxytocin have been found to signal relationship distress in women. Whether vasopressin underlies additional aspects of men's affiliative responses to stress is currently unknown, but this issue merits additional exploration.

Social Issue Implications of Tend and Befriend

What critical social issues and problems are raised or resolved by insights generated by the tend and befriend theory? One important social issue on which the theory helps to shed light is sex differences in life expectancy. Men are especially vulnerable to early mortality due to homicide, suicide, coronary heart disease, and disorders related to substance abuse for coping with stress. Women, on the other hand, enjoy a substantial advantage in mortality in most countries of the world. Only countries in which women are denied access to healthcare or those in which deaths during childbirth are still common show a reverse gap. Of interest is the fact that the causes of death that largely account for men's early mortality are those related to the fight or flight response, namely aggressive responses to stress, withdrawal in the form of substance abuse, and coronary artery disease, the risk for which is exacerbated by frequent or recurrent stress exposure. By contrast, women more reliably turn to their social contacts in times of stress, responses that are, as just noted, protective of health and longevity. The fact that men may be somewhat more likely to cope with stress via fight or flight and women to cope with stress via tend and befriend may help to explain the worldwide gender gap in mortality.

Building on this point, the theory may help to explain sudden increases in mortality rates [p. 43 ↓] that are seen in countries experiencing substantial economic and political turmoil. For example, following the end of the Soviet Union in 1989, destabilization of the Eastern European social environment left many people socially unattached. Whereas unattached women often came together in informal groups with other women and children and shared the management of tasks of daily life, men more often coped with the same instability through alcohol abuse, smoking, and aggressive encounters with other males (Bobak and Marmot, 1996; Stone, 2000). Men in the former Eastern

European nations subsequently experienced an abrupt decline of approximately seven years in life expectancy in a mere five years, worse than that sustained during World War II. This decline in life expectancy was explained substantially by deaths among unattached men (Bobak and Marmot, 1996; Stone, 2000). Having a theoretical basis for explaining events such as these, and the psychological and biological mechanisms that underpin them, constitutes an advance that may help governments to anticipate similar problems and intervene, so that similar upheaval does not result in similar carnage.

The theory and the evidence consistent with tend and befriend also point to the importance of making affiliative opportunities available to people when they are under stress. Affiliation is inherently comforting, even when no explicit efforts at social support are elicited or provided. Although questions have recently been raised about the necessity or value of making psychological counseling available to people in the immediate aftermath of a trauma or highly stressful event, providing people with opportunities for companionship which they may utilize in whatever ways are most comforting may be useful interventions.

A final issue on which the tend and befriend model sheds light concerns the relative benefits and costs of altruistic behavior. Conceptualizations of altruism and social support have been guided by the implicit assumption that support is beneficial for the recipient but costly for the provider. This viewpoint has been shaped by the evolutionary perspective on altruism which addresses the paradox: how do we pass on our altruistic genes to future generations if those genes put us at potential risk? That is, when one person helps another in times of threat, the likelihood that the helper will be harmed can be high. Research on the physical and psychological costs of caregiving would seem to support the position that altruism is inherently costly. From an evolutionary perspective, altruism is largely rescued by the concept of reciprocal altruism (Hamilton, 1963; Trivers, 1971).

Evidence consistent with the tend and befriend model, however, demonstrates that the nurturant behavior of tending not only benefits the offspring but also benefits the tender. That is, tending behavior following a stressful encounter not only downregulates the stress systems of offspring, but also the stress systems of the mother; thus, to the extent that caregivers are providing nurturant behavior to others, their own stress systems may be benefited in cumulative fashion. Research by Brown et al.

(2003a) found that death rates were significantly lower among people who provided instrumental support to friends, relatives and neighbors, and emotional support to their spouses. Receiving support did not affect mortality once giving support was statistically controlled. Thus, this study provides important evidence that the giving of support can promote health or retard illness progression. There are psychological benefits of giving support to others as well. Giving support may cement personal relationships, provide a sense of meaning or purpose, and signify that one matters, all of which have been found to promote wellbeing.

It is likely, then, that the benefits of providing support and the apparent absence of anticipated costs may work through some of the same physiological and neuroendocrine pathways whereby the receipt of support or perception of it from others achieves its benefits. In addition, the anxiolytic properties of oxytocin, coupled with its established [p. 44 ↓] role in downregulating SNS and HPA axis responses to stress, may help to provide an understanding of the mental and physical health benefits of providing social support as well as receiving it.

Thoughts for the Future

Useful theories in social psychology are marked not only by their ability to explain a body of data and generate specific hypotheses, but also by their ability to expand and grow as new evidence emerges. The tend and befriend theory is no different. Among the extensions that have occurred since the original publication of the theory in 2000 (Taylor et al., 2000) is the accumulating evidence for a biologically based signaling system that may alert people to the need to enhance their social contacts and impel them to do so as well; the fact that plasma oxytocin in women and plasma vasopressin in men are elevated in conjunction with breaches in social contact is a relatively recent discovery.

A second extension involves the integration of evidence concerning the biological stress regulatory effects of maternal nurturance into the theory. Although insights regarding the importance of maternal nurturance for developing offspring's biological stress regulatory systems have been accumulating, the mechanisms underlying these effects remained unknown until relatively recently. Both animal and human evidence now points to some

of the biological mechanisms, including epigenetic mechanisms, that help to explain why maternal nurturance in times of stress is so vital to the development of offspring's biological systems and socio-emotional skills for managing stress. Integrating this evidence more fully with the literature on the biological and psychological bases of attachment is an important future step.

A third advance concerns potential genetic contributions to the processes detailed in the tend and befriend theory. At the time the theory was developed, little was known about what specific genes might contribute to tending, befriending, and social behavior more generally. The only evidence for genetic contributions to social behavior was from twin studies indicating a large genetic contribution to the experience of social support (Kessler et al., 1992). Research has now enabled the identification of specific genes within the opioid, oxytocin, vasopressin, dopamine, and serotonin systems that may be implicated in the processes detailed here (see Way and Taylor, 2011, for a review).

Just as accumulating evidence has expanded the theory in certain ways, so has additional research identified certain problems to be resolved. For example, elevated plasma oxytocin is tied to social distress, whereas exogenous administration of oxytocin is tied to a sense of calm and relaxation. This paradox has yet to be fully resolved. As noted, biological underpinnings of men's social responses to stress have yet to be as rigorously explored, although some recent progress has been made (Taylor et al., 2009). No doubt the future will pose additional challenges for the theory.

Tend and Befriend: Biobehavioral Redundancy?

Many of the benefits of social contact appear to result from sheer proximity and not necessarily from the socially supportive transactions that have typically been studied by social psychologists and health psychologists. For example, social ties are consistently found to be mentally and physically health protective in both stressful and non-stressful environments, whereas social support transactions appear to be most beneficial in situations of stress. Qualifying this last finding further, a large number of circumstances have been identified in which social support transactions are unsuccessful or have

unintended negative consequences. Bolger and colleagues (e.g., Bolger and Amarel, 2007), for example, have suggested that [p. 45 ↓] invisible support (i.e., support provided by a person without the recipient's awareness) is more beneficial to emotional functioning than social support efforts that are recognized by both the giver and the recipient as intended. Awareness that one is being supported may represent a threat to self-esteem. Research on the matching hypothesis, namely the idea that social support is most effective when it matches the need of the recipient (Cohen and McKay, 1984), also indicates a variety of circumstances under which mismatches between the type of support delivered or the person delivering it exacerbate stress (see Taylor, 2009, for a review). Social contacts during stressful times have the potential to be negative, and research has shown that negative interactions can have a worse effect on mental and physical health functioning than positive effects achieve beneficial effects (e.g., Rook, 1984). Other misfired efforts at social support have also been identified (see Taylor, 2009, for a review).

These findings suggest that simple proximity and the perception of support may be especially beneficial but not necessarily its use. That is, much of the benefit of social support may come from the perception that it is available and not necessarily its actual engagement (Thoits, 1995; Taylor, 2009). A biobehavioral theory of affiliation, as tend and befriend is, has little trouble with this paradox. Either the psychological benefits associated with social contact, the hormonal underpinnings of social contact, or both, may produce many of the beneficial biological and psychological consequences that consistently predict wellbeing both in stressful and nonstressful times.

A common observation in biological research concerns the redundancy that exists in the human being for sustaining vital biological functions. For example, there are five different mechanisms that ensure that the stomach can digest food. Similarly, people are endowed with two eyes, two hands, two legs, two kidneys, and the like. Not all vital functions are backed up through redundancy of course, the heart being an obvious counterexample. However, many are, and it may be useful to think about psychological processes as contributing to the human beings' redundancy to ensure vital processes.

At the outset of this chapter, the point was made that social living in general and affiliation in response to threat in particular are essential to human survival. It is unlikely that these critical responses would be left to chance or even to a single underlying

psychological or biological mechanism. Rather, the fact that there is both biological and psychological evidence for tending and befriending in response to stress suggests that these may be interrelated but semi-redundant pathways that ensure that social responses to threat take place, so as to protect human beings and ensure their survival.

Although a theory that focuses on affiliative responses to threat, as the tend and befriend theory does, stresses the important protective role that these processes have, it must be noted that tending and befriending also promote human growth as well. Through affiliation with others in times of stress and the ability to draw on mental representations of relationships, people acquire the resources to explore and grow both emotionally and intellectually in environments that ensure social connection.

Acknowledgment

Preparation of this manuscript was supported by a grant from NSF (SES-0525713).

References

- Allen, K., Blascovich, J. and Mendes, W.B. (2002) Cardiovascular reactivity and the presence of pets, friends, and spouses: The truth about cats and dogs . *Psychosomatic Medicine* , 64, 727–739. <http://dx.doi.org/10.1097/01.PSY.0000024236.11538.41>
- Altemus, M.P., Deuster, A., Galliven, E., Carter, C.S. and Gold, P.W. (1995) Suppression of hypothalamic-pituitary-adrenal axis response to stress in lactating women . *Journal of Clinical Endocrinology and Metabolism* , 80, 2954–2959. <http://dx.doi.org/10.1210/jc.80.10.2954>
- Anckarsäter, H. and Cloninger, CR. (2007) The genetics of empathy and its disorders . In T. Farrow, ed. and P. Woodruff (eds), *Empathy in Mental Illness* , pp. 261–288. New York: Cambridge University Press.

Berkman, L.F. and Syme, S.L. (1979) Social networks, host resistance, and mortality: A nine-year follow up study of Alameda County residents . American Journal of Epidemiology , 109, 186–204.

Bobak, M. and Marmot, M. (1996) East-West mortality divide and its potential explanations: Proposed research agenda . British Medical Journal , 312, 421–425.
<http://dx.doi.org/10.1136/bmj.312.7028.421>

Bolger, N. and Amarel, D. (2007) Effects of support visibility on adjustment to stress: Experimental evidence . Journal of Personality and Social Psychology , 92, 458–475.
<http://dx.doi.org/10.1037/0022-3514.92.3.458>

Bowlby, J. (1982) Attachment and Loss, Vol. 1, Attachment (2nd edition). New York: Basic Books. (Original work published 1969).

Brown, S.L., Nesse, R.M., Vinokur, A.D. and Smith, D.M. (2003a) Providing social support may be more beneficial than receiving it: Results from a prospective study of mortality . Psychological Science , 14, 320–327. <http://dx.doi.org/10.1111/1467-9280.14461>

Brown, J.L., Sheffield, D., Leary, M.R. and Robinson, M.E. (2003b) Social support and experimental pain . Psychosomatic Medicine , 65, 276–283. <http://dx.doi.org/10.1097/01.PSY.0000030388.62434.46>

Cacioppo, J.T. and Hawkley, L.C. (2003) Social isolation and health, with an emphasis on underlying mechanism . Perspectives in Biology and Medicine , 46, S39–S52.

Cannon, W.B. (1932) The Wisdom of the Body . New York: Norton.

Carlson, M. and Earls, F. (1997) Psychological and neuroendocrinological sequelae of early social deprivation in institutionalized children in Romania . Annals of the New York Academy of Sciences , 807, 419–428. <http://dx.doi.org/10.1111/j.1749-6632.1997.tb51936.x>

Carter, CS. (1998) Neuroendocrine perspectives on social attachment and love . *Psychoneuro-endocrinology* , 23, 779–818. <http://dx.doi.org/10.1016/S0306-4530%2898%2900055-9>

Carter, C.S., Lederhendler, I.I. and Kirkpatrick, B. (1999) *The Integrative Neurobiology of Affiliation* . Cambridge, MA: MIT Press.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I.W., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A. and Poulton, R. (2003) Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene . *Science* , 301, 386–389. <http://dx.doi.org/10.1126/science.1083968>

Champagne, F.A., Chretien, P., Stevenson, C.W., Zhang, T.Y., Gratton, A. and Meaney, M.J. (2004) Variations in nucleus accumbens dopamine associated with individual differences in maternal behavior in the rat . *Journal of Neuroscience* , 24, 4113–4123. <http://dx.doi.org/10.1523/JNEUROSCI.5322-03.2004>

Chiadra, P. and Legros, J.J. (1981) Intravenous injection of synthetic oxytocin induces a decrease of cortisol plasma level in normal man . *Comptes rendus des seances de la Societe de biologie et de ses filiales* , 175, 546–954.

Cohen, S. and McKay, G. (1984) Social support, stress, and the buffering hypothesis: A theoretical analysis . In A. Baum, ed. , S.E. Taylor, ed. , and J. Singer (eds), *Handbook of Psychology and Health* , 4, 253–268. Hillsdale, NJ: Erlbaum.

Cohen, S.D., Sharma, T., Acquaviva, K., Peterson, R.A., Patel, S.S. and Kimmel, P.L. (2007) Social support and chronic kidney disease: An update . *Advances in Chronic Kidney Disease* , 14, 335–344. <http://dx.doi.org/10.1053/j.ackd.2007.04.007>

Collins, N.L., Dunkel-Schetter, C., Lobel, M. and Scrimshaw, S.C.M. (1993) Social support in pregnancy: Psychosocial correlates of birth outcomes and post-partum depression . *Journal of Personality and Social Psychology* , 65, 1243–1258. <http://dx.doi.org/10.1037/0022-3514.65.6.1243>

Collins, N.L. and Feeney, B.C. (2000) A safe haven: An attachment theory perspective on support seeking and caregiving in intimate relationships . *Journal of Personality and Social Psychology* , 78, 1053–1073. <http://dx.doi.org/10.1037/0022-3514.78.6.1053>

Detillion, C.E., Craft, T.K., Glasper, E.R., Prendergast, B.J. and DeVries, C. (2004) Social facilitation of wound healing . *Psychoneuroendocrinology* , 29, 1004–1011. <http://dx.doi.org/10.1016/j.psyneuen.2003.10.003>

Eisenberger, N.I., Lieberman, M.D. and Williams, K.D. (2003) Does rejection hurt? An fMRI study of social exclusion . *Science* , 302, 290–292. <http://dx.doi.org/10.1126/science.1089134>

Fleming, R., Baum, A., Gisriel, M.M. and Gatchel, R.J. (1982) Mediating influences of social support on stress at Three Mile Island . *Journal of Human Stress* , 8, 14–22. <http://dx.doi.org/10.1080/0097840X.1982.9936110>

Fraley, R.C (2002) Attachment stability from infancy to adulthood: Meta-analysis and dynamic modeling of developmental mechanisms . *Personality and Social Psychology Review* , 6, 123–151. http://dx.doi.org/10.1207/S15327957PSPR0602_03

Francis, D., Diorio, J., Liu, D. and Meaney, M.J. (1999) Nongenomic transmission across generations of maternal behavior and stress responses in the rat . *Science* , 286, 1155–1158. <http://dx.doi.org/10.1126/science.286.5442.1155>

Grippo, A.J., Gerena, D., Huang, J., Kumar, N., Shah, M., Ughreja, R. and Carter, CS. (2007) Social isolation induces behavioral and neuroendocrine disturbances relevant to depression in female and male prairie voles . *Psychoneuroendocrinology* , 32, 966–980. <http://dx.doi.org/10.1016/j.psyneuen.2007.07.004>

Gunnar, M.R., Brodersen, L., Krueger, K. and Rigatuso, J. (1996) Dampening of adrenocortical responses during infancy: Normative changes and individual differences . *Child Development* , 67, 877–889. <http://dx.doi.org/10.2307/1131867>

Hamilton, W.D. (1963) The evolution of altruistic behavior . *The American Naturalist* , 97, 354–356. <http://dx.doi.org/10.1086/497114>

- Hart, J., Gunnar, M. and Cicchetti, D. (1996) Altered neuroendocrine activity in maltreated children related to symptoms of depression . Development and Psychopathology , 8, 201–214. <http://dx.doi.org/10.1017/S0954579400007045>
- Heinrichs, M., Baumgartner, T., Kirshbaum, C. and Ehlert, U. (2003) Social support and oxytocin interact to suppress Cortisol and subjective responses to psychological stress . Biological Psychiatry , 54, 1389–1398. <http://dx.doi.org/10.1016/S0006-3223%2803%2900465-7>
- Holahan, C.J., Moos, R.H., Holahan, C.K. and Brennan, P.L. (1997) Social context, coping strategies, and depressive symptoms: An expanded model with cardiac patients . Journal of Personality and Social Psychology , 72, 918–928. <http://dx.doi.org/10.1037/0022-3514.72.4.918>
- House, J.S., Landis, K.R. and Umberson, D. (1988) Social relationships and health . Science , 241, 540–545. <http://dx.doi.org/10.1126/science.3399889>
- Insel, T.R. (1997) A neurobiological basis of social attachment . American Journal of Psychiatry , 154, 726–735.
- Jalowiec, J.E., Calcagnetti, D.J. and Fanselow, M.S. (1989) Suppression of juvenile social behavior requires antagonism of central opioid systems . Pharmacology Biochemistry and Behavior , 33, 697–700. <http://dx.doi.org/10.1016/0091-3057%2889%2990409-7>
- Jamner, L.D., Alberts, J., Leigh, H. and Klein, L.C. (1998) Affiliative need and endogenous opioids. Paper presented to the Society of Behavioral Medicine annual meeting , March 1998 New Orleans, LA.
- Kaufman, J., Yang, B.Z., Douglas-Palumberi, H., Houshyar, S., Lipschitz, D., Krystal, J.H. and Gelernter, J. (2004) Social supports and serotonin transporter gene moderate depression in maltreated children . Proceedings of the National Academy of Sciences USA , 101, 17316–17321. <http://dx.doi.org/10.1073/pnas.0404376101>
- Kendrick, K.M. and Keverne, E.B. (1989) Effects of intracerebroventricular infusions of naltrexone and phentolamine on central and peripheral oxytocin release and on

maternal behaviour induced by vaginocervical stimulation in the ewe . Brain Research , 505, 329–332. <http://dx.doi.org/10.1016/0006-8993%2889%2991462-5>

Kessler, R.C., Kendler, K.S., Heath, A.C., Neale, M.C. and Eaves, L.J. (1992) Social support, depressed mood, and adjustment to stress: A genetic epidemiological investigation . Journal of Personality and Social Psychology , 62, 257–272. <http://dx.doi.org/10.1037/0022-3514.62.2.257>

King, K.B., Reis, H.J., Porter, L.A. and Norsen, L.H. (1993) Social support and long-term recovery from coronary artery surgery: Effects on patients and spouses . Health Psychology , 12, 56–63. <http://dx.doi.org/10.1037/0278-6133.12.1.56>

Knafo, A., Israel, S., Darvasi, A., Bachner-Melman, R., Uzefovsky, F., Cohen, L. et al. (2008) Individual differences in allocation of funds in the dictator game associated with length of the arginine vasopressin 1a receptor RS3 promoter region and correlation between RS3 length and hippocampal mRNA . Genes, Brain, and Behavior , 7, 266–275. <http://dx.doi.org/10.1111/j.1601-183X.2007.00341.x>

Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U. and Fehr, E. (2005) Oxytocin increases trust in humans . Nature , 435, 673–676. <http://dx.doi.org/10.1038/nature03701>

Kulik, J.A. and Mahler, H.I.M. (1993) Emotional support as a moderator of adjustment and compliance after coronary artery bypass surgery: A longitudinal study . Journal of Behavioral Medicine , 16, 45–64. <http://dx.doi.org/10.1007/BF00844754>

Levine, S. and Wiener, S.G. (1988) Psychoendocrine aspects of mother-infant relationships in nonhuman primates . Psychoneuroimmunology 13, 143–154.

Light, K.C., Smith, T.E., Johns, J.M., Brownley, K.A., Hofheimer, J.A. and Amico, J.A. (2000) Oxytocin responsivity in mothers of infants: A preliminary study of relationships with blood pressure during laboratory stress and normal ambulatory activity . Health Psychology , 19, 560–567. <http://dx.doi.org/10.1037/0278-6133.19.6.560>

Lim, M.M., Wang, Z., Olazabal, D.E., Ren, X., Terwilliger, E.F. and Young, L.J. (2004) Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene . *Nature* , 429, 754–757. <http://dx.doi.org/10.1038/nature02539>

Lin, N., Ye, X. and Ensel, W. (1999) Social support and depressed mood: A structural analysis . *Journal of Health and Social Behavior* , 40, 344–359. <http://dx.doi.org/10.2307/2676330>

Liu, D., Diorio, J., Tannenbaum, B., Caldji, C., Francis, D., Freedman, A., Sharma, S., Pearson, D., Plotsky, P.M. and Meaney, M.J. (1997) Maternal care, hippocampal glucocorticoid receptors, and hypothalamic–pituitary–adrenal responses to stress . *Science* , 277, 1659–1662. <http://dx.doi.org/10.1126/science.277.5332.1659>

Mantella, R.C., Vollmer, R.R., Li, X. and Amico, J.A. (2003) Female oxytocin-deficient mice display enhanced anxiety-related behavior . *Endocrinology* , 144, 2291–2296. <http://dx.doi.org/10.1210/en.2002-0197>

Marteau, T.M., Bloch, S. and Baum, J.D. (1987) Family life and diabetic control . *Journal of Child Psychology and Psychiatry* , 28, 823–833. <http://dx.doi.org/10.1111/j.1469-7610.1987.tb00671.x>

Martel, F.L., Nevison, C.M., Rayment, F.D., Simpson, M.J.A. and Keverne, E.B. (1993) Opioid receptor blockade reduces maternal affect and social grooming in rhesus monkeys . *Psychoneuroimmunology* , 75, 307–321.

McCarthy, M.M. (1995) Estrogen modulation of oxytocin and its relation to behavior . *Advances in Experimental Medicine and Biology* , 395, 235–245.

McEwen, B.S. (1998) Protective and damaging effects of stress mediators . *New England Journal of Medicine* , 338, 171–179. <http://dx.doi.org/10.1056/NEJM199801153380307>

Meaney, M.J. (2001) Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations . *Annual Review of Neuroscience* , 24, 1161–1192. <http://dx.doi.org/10.1146/annurev.neuro.24.1.1161>

Meaney, M.J. and Szyf, M. (2005) Environmental programming of stress responses through DNA methylation: life at the interface between a dynamic environment and a fixed genome . *Dialogues in Clinical Neurosciences* , 7, 103–123.

Modahl, C. and Newton, N. (1979) Mood state differences between breast and bottle-feeding mothers . *Emotion and Reproduction, Proceedings of the Serono Symposia* , 20B, 819–822.

Moles, A., Kieffer, B.L. and D'Amato, F.R. (2004) Deficit in attachment behavior in mice lacking the μ -opioid receptor gene . *Science* , 304, 1983–1985. <http://dx.doi.org/10.1126/science.1095943>

Nachmias, M., Gunnar, M.R., Mangelsdorf, S., Parritz, R.H. and Buss, K. (1996) Behavioral inhibition and stress reactivity: The moderating role of attachment security . *Child Development* , 67, 508–522. <http://dx.doi.org/10.2307/1131829>

Panksepp, J. (1998) *Affective Neuroscience* . New York: Oxford University Press.

Panksepp, J., Nelson, E. and Bekkedal, M. (1999) Brain systems for the mediation of social separation distress and social-reward . In C.S. Carter, ed. , I.I. Lederhendler, ed. , and B. Kirkpatrick (eds), *The Integrative Neurobiology of Affiliation* , pp. 221–244. Cambridge, MA: MIT Press.

Penninx, B.W.J.H., van Tilburg, T., Boeke, A.J.P., Deeg, D.J.H., Kriegsman, D.M.W. and van Eijk, J.T.M. (1998) Effects of social support and personal coping resources on depressive symptoms: Different for various chronic diseases? *Health Psychology* , 17, 551–558. <http://dx.doi.org/10.1037/0278-6133.17.6.551>

Petersson, M., Alster, P., Lundeberg, T. and Uvnäs-Moberg, K. (1996) Oxytocin causes a long-term decrease of blood pressure in female and male rats . *Physiology and Behavior* , 60, 1311–1315. <http://dx.doi.org/10.1016/S0031-9384%2896%2900261-2>

Repetti, R.L., Taylor, S.E. and Saxbe, D. (2007) The influence of early socialization experiences on the development of biological systems . In J. Grusec, ed. and P. Hastings (ed.), *Handbook of Socialization* (pp. 124–152). New York: Guilford.

Repetti, R.L., Taylor, S.E. and Seeman, T.E. (2002) Risky families: Family social environments and the mental and physical health of offspring . *Psychological Bulletin* , 128, 330–366. <http://dx.doi.org/10.1037/0033-2909.128.2.330>

Rook, K.S. (1984) The negative side of social interaction: Impact on psychological well-being . *Journal of Personality and Social Psychology* , 46, 1097–1108. <http://dx.doi.org/10.1037/0022-3514.46.5.1097>

Rutledge, T., Reis, S.E., Olson, M., Owens, J., Kelsey, S.F., Pepine, C.J. et al. (2004) Social networks are associated with lower mortality rates among women with suspected coronary disease: The National Heart, Lung, and Blood Institute-sponsored women's ischemia syndrome evaluation study . *Psychosomatic Medicine* , 66, 882–888. <http://dx.doi.org/10.1097/01.psy.0000145819.94041.52>

Sarason, B.R., Sarason, I.G. and Gurung, R.A.R. (1997) Close personal relationships and health outcomes: A key to the role of social support . In S. Duck (ed.) *Handbook of Personal Relationships* , pp. 547–573. New York: Wiley.

Seeman, T.E., Lusignolo, T.M., Albert, M. and Berkman, L. (2001) Social relationships, social support, and patterns of cognitive aging in healthy, high-functioning older adults: MacArthur Studies of Successful Aging . *Health Psychology* , 20, 243–255. <http://dx.doi.org/10.1037/0278-6133.20.4.243>

Simeon, D., Greenberg, J., Nelson, D., Schmeider, J. and Hollander, E. (2005) Dissociation and posttraumatic stress 1 year after the World Trade Center disaster: Follow-up of a longitudinal study . *Journal of Clinical Psychiatry* , 66, 231–237. <http://dx.doi.org/10.4088/JCP.v66n0212>

Sorkin, D., Rook, K.S. and Lu, J.L. (2002) Loneliness, lack of emotional support, lack of companionship, and the likelihood of having a heart condition in an elderly sample . *Annals of Behavioral Medicine* , 24, 290–298. http://dx.doi.org/10.1207/S15324796ABM2404_05

Spitz, R.A. and Wolff, K.M. (1946) Anaclitic depression: An inquiry into the genesis of psychiatric conditions in early childhood, II . In A. Freud et al. (eds), *The Psychoanalytic Study of the Child* , 2, 313–342. New York: International Universities Press.

Stone, A.A., Mezzacappa, E.S., Donatone, B.A. and Gonder, M. (1999) Psychosocial stress and social support are associated with prostate-specific antigen levels in men: Results from a community screening program . *Health Psychology* , 18, 482–486. <http://dx.doi.org/10.1037/0278-6133.18.5.482>

Stone, R. (2000) Stress: The invisible hand in eastern Europe's death rates . *Science* , 288, 1732–1733. <http://dx.doi.org/10.1126/science.288.5472.1732>

Suomi, S.J. (1987) Genetic and maternal contributions to individual differences in rhesus monkey biobehavioral development . In N.A. Krasnagor, ed. , E.M. Blass, ed. , MA. Hofer, ed. , and W.P. Smotherman (eds), *Perinatal Development: A Psychobiological Perspective* , pp. 397–420. New York: Academic Press.

Suomi, S.J. (1999) Attachment in rhesus monkeys . In J. Cassidy, ed. and P. Shaver (eds), *Handbook of Attachment: Theory, Research, and Clinical Applications* , pp. 181–197. New York: Guilford Press.

Tamres, L., Janicki, D. and Helgeson, V.S. (2002) Sex differences in coping behavior: A meta-analytic review . *Personality and Social Psychology Review* , 6, 2–30. http://dx.doi.org/10.1207/S15327957PSPR0601_1

Taylor, S.E. (2002) *The Tending Instinct: How Nurturing Is Essential to Who We Are and How We Live* . New York: Holt.

Taylor, S.E. (2009) Social support: A review . In H.S. Friedman (ed.), *Oxford Handbook of Health Psychology* . New York: Oxford University Press.

Taylor, S.E., Gonzaga, G., Klein, L.C., Hu, P., Greendale, G.A. and Seeman, T.E. (2006a) Relation of oxytocin to psychological stress responses and hypothalamic-pituitary-adrenocortical axis activity in older women . *Psychosomatic Medicine* , 68, 238–245. <http://dx.doi.org/10.1097/01.psy.0000203242.95990.74>

Taylor, S.E., Klein, L.C., Lewis, B.P., Gruenewald, T.L., Gurung, R.A.R., and Updegraff, J.A. (2000) Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight . *Psychological Review* , 107, 411–429. <http://dx.doi.org/10.1037/0033-295X.107.3.411>

Taylor, S.E., Saphire-Bernstein, S. and Seeman, T.E. (2009) Plasma oxytocin in women and plasma vasopressin in men are markers of distress in primary relationships . Psychological Science , 21, 3–7. <http://dx.doi.org/10.1177/0956797609356507>

Taylor, S.E., Way, B.M., Welch, W.T., Hilmert, C.J., Lehman, B.J. and Eisenberger, N.I. (2006b) Early family environment, current adversity, the serotonin transporter promoter polymorphism, and depressive symptomatology . Biological Psychiatry , 60, 671–676. <http://dx.doi.org/10.1016/j.biopsych.2006.04.019>

Thoits, P.A. (1995) Stress, coping and social support processes: Where are we? What next? Journal of Health and Social Behavior , (Extra Issue), 53–79. <http://dx.doi.org/10.2307/2626957>

Trivers, R.L (1971) The evolution of reciprocal altruism . The Quarterly Review of Biology , 46, 35–57. <http://dx.doi.org/10.1086/406755>

Turner, R.A., Altemus, M., Enos, T., Cooper, B. and McGuinness, T. (1999) Preliminary research on plasma oxytocin in normal cycling women: Investigating emotion and interpersonal distress . Psychiatry , 62, 97–113.

Turner-Cobb, J.M., Gore-Felton, C., Marouf, F., Koopman, C., Kim, P., Israelski, D. and Spiegel, D. (2002) Coping, social support, and attachment style as psychosocial correlates of adjustment in men and women with HIV/AIDS . Journal of Behavioral Medicine , 25, 337–353. <http://dx.doi.org/10.1023/A:1015814431481>

Uvnäs-Moberg, K. (1997) Oxytocin linked antistress effects – the relaxation and growth response . Acta Psychologica Scandinavica , 640, 38–42.

Uvnäs-Moberg, K., Ahlenius, S., Hillegaart, V. and Alster, P. (1994) High doses of oxytocin cause sedation and low doses cause an anxiolytic-like effect in male rats . Pharmacology, Biochemistry, and Behavior , 49, 101–106.

VanderPlate, C., Aral, S.O. and Magder, L. (1988) The relationship among genital herpes simplex virus, stress, and social support . Health Psychology , 7, 159–168. <http://dx.doi.org/10.1037/0278-6133.7.2.159>

Way, B.M. and Taylor, S.E. (2011) Genetic factors in social pain . In G. MacDonald, ed. and L.A. Jensen-Campbell (eds), Social Pain: A Neuroscientific, Social, Clinical, and Developmental Analysis , pp. 95–119. Washington, DC: American Psychological Association.

Wiklund, I., Oden, A., Sanne, H., Ulvenstam, G., Wilhemsson, C. and Wilhemsen, L. (1988). Prognostic importance of somatic and psychosocial variables after a first myocardial infarction . American Journal of Epidemiology , 128, 786–795.

Wills, T.A. (1991) Social support and interpersonal relationships . In M.S. Clark (ed.), Prosocial Behavior , pp. 265–289. Newbury Park, CA: Sage.

Zak, P.J., Kurzban, R. and Matzner, W.T. (2004) The neurobiology of trust . Annals of the New York Academy of Sciences , 1032, 224–221. <http://dx.doi.org/10.1196/annals.1314.025>

<http://dx.doi.org/10.4135/9781446249215.n3>

Handbook of Theories of Social Psychology: Volume 1

The Evaluative Space Model

Contributors: John T. Cacioppo & Gary G. Berntson & Catherine J. Norris & Jackie K. Gollan

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "The Evaluative Space Model"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n4>

Print pages: 50-age73

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n4>

[p. 50 ↓]

Chapter 3: The Evaluative Space Model

John T. Cacioppo, ed. , Gary G. Berntson, ed. , Catherine J. Norris, ed. , and Jackie K. Gollan, ed.

Abstract

The ability to differentiate hostile from hospitable stimuli is ubiquitous in animals. The evaluative space model (ESM) is a theory of the functional structure and operating characteristics of these evaluative processes across levels of the neuraxis, ranging from spinal cord reflexes to the executive functions of the frontal lobes (e.g., impulse control). According to the ESM, physical constraints limit behavioral expressions and incline behavioral predispositions toward a bipolar organization, but this bipolar organization is posited to be the consequence of multiple operations, including motivational activation function for positivity (appetition) and the activation function negativity (aversion). The partial segregation of positive and negative evaluative processes afforded evolution that opportunity to sculpt distinct activation functions for positivity and negativity, and permits greater flexibility of these evaluative processes such as reciprocal activation, uncoupled activation, or coactivation/coinhibition. The result is a much more flexible and adaptable affect system of evaluative processes than would be provided were evaluative processes characterized simply as a bipolar (positive-negative) activation function.

The Evaluative Space Model

The ABCs of the mind have been described as affect, behavior, and cognition. These were once thought to represent independent sets of psychological structures and processes, but this is no longer the case. Theory and research on emotional contagion, for instance, underscore the interplay between affect and behavior; work on embodied cognition stresses the overlap between behavior and cognition; studies of motivated attention and cognition focus on the interchange between affect and cognition; and research on attitudes emphasizes the interplay among all three. Nevertheless, there are sufficient distinctions among affect, behavior, and cognition that psychological theories of each are needed. Of the ABCs of the mind, the least well developed theoretically is the topic of affect.

The term “affect” in human studies has been used to refer to feelings beyond those of [p. 51 ↓] the traditional senses, with an emphasis on the experience of emotions and variations in hedonic tone. Accordingly, the scientific study of human affect and emotion has tended to emphasize reportable feeling states. Studies of the conceptual organization of affect and emotion indicate that people represent feelings and emotions in terms of a circular order around the perimeter of the space defined by a bipolar valence dimension and an orthogonal dimension labeled activation (i.e., a circumplex; for example, see Russell, 1980) or, alternatively, by a space defined by two bipolar valence dimensions, one ranging from low positive affect to high negative affect and a second ranging from low negative affect to high positive affect (Green et al., 1999; Watson and Tellegen, 1985). As LeDoux noted, however:

It is widely recognized that most cognitive processes occur unconsciously, with only the end products reaching awareness, and then only sometimes. Emotion researchers, though, did not make this conceptual leap. They remained focused on subjective emotional experience ... The main lesson to be learned ... is that emotion researchers need to figure out how to escape from the shackles of subjectivity if emotion research is to thrive (LeDoux, 2000: 156).

There is an understandable appeal to settling for feelings as the appropriate data to model in the area of affect. It is these feelings that some theorists seek to describe, understand, and explain. The structure and processes underlying mental contents are not readily apparent, however, and most cognitive processes occur unconsciously with only selected outcomes reaching awareness. Over millions of years of evolution, efficient and manifold mechanisms have evolved for differentiating hostile from hospitable stimuli and for organizing adaptive responses to these stimuli. These are critically important functions for the evolution of mammals, and the integrated set of mechanisms that serve these functions can be thought of as an “affect system.” It is this affect system – its architecture and operating characteristics – that is the focus of the evaluative space model (ESM) (Cacioppo and Berntson, 1994, 1999; Cacioppo et al., 1997, 1999).

We use the term “evaluative” rather than “affective” in the name of our theory to emphasize the focus on the affect system as the production mechanism for discrimination, valuation, feelings, and adaptive behavior. On the stimulus side, evaluative processes refer to the discrimination of a stimulus – or features of a stimulus – as appetitive or aversive, hostile or hospitable, pleasant or unpleasant, threatening or nurturing. Such discriminations, according to the ESM, extend beyond what can be verbalized, as spinal cord reflexes and affective priming amply demonstrate. On the response side, evaluative processes organize behavioral responses to promote appropriate approach or withdrawal, advancement or retreat, movement toward or away, attack or avoidance, nurturance or defense, acceptance or rejection. The evaluative operations intervening between these stimulus and response elements are also part of the affect system.

The ESM is a general theoretical formulation of the functional architecture and operating characteristics of the affect system, the postulates of which are summarized in [Table 3.1](#). For instance, theorists have argued about whether the affect system should be construed in terms of discrete emotions or abstract dimensional structures. According to the ESM, these are not mutually exclusive but rather each is a theoretical representation at a specific level of organization (*level of analysis postulate*). There are important distinctions among the positive emotions and among the negative emotions, but the positive emotions are generally more similar to each other than they are to the negative emotions and vice versa. The ESM posits a super-ordinate dimensional

structure representing appetitive predispositions, positive affects, and emotions (termed positivity), and a superordinate dimensional structure representing defensive predispositions, negative affects, and emotions (termed negativity). If appetitive predispositions and positive [p. 52 ↓] [p. 53 ↓] [p. 54 ↓] affects and emotions were stars, the positivity dimension might be thought of as the galaxy constituted by these stars, and if defensive predispositions, negative affects, and emotions were stars, the negativity dimension could be conceived as a different galaxy comprising this very different cluster of stars. That is, these positive and negative dimensions are not treated as equivalent in their constitution, operations, or consequences (*functional separability postulate*; see [Table 3.1](#)).

Postulate	Definition	Additional detail
Level of analysis	There are distinctions among both positive and negative emotions, but positive emotions are more similar to each other than they are to negative emotions, and vice versa.	A single valence continuum does not capture the structure and operating characteristics of affect system.
Functional separability	Positivity and negativity are not equivalent in their constitution, operations, or consequences.	There is a superordinate dimensional structure representing appetitive predispositions, positive affects, and emotions, as well a superordinate dimensional structure representing defensive predispositions, negative affects, and emotions.
Heteroscedacity	The constellation of antecedents, emotions, expressions, and response is more diverse for negativity than positivity.	When an event elicits a positive emotion, staying the course is sufficient; when negative emotion is elicited, an adaptive

		response may vary greatly across eliciting events.
Energetic efficiency	Behavior in future encounters with target stimuli will tend to be more expected and stable when organized in terms of a bipolar evaluative dimension.	Behavioral and cognitive efficiency and a reduction in stress is served by mental representations of general action predispositions toward classes of stimuli.
Evaluative activation	Affect is a joint function of positively and negatively valent activation functions.	The resultant output disposition(s) is not necessarily the simple arithmetic mean of the separate functions. The net disposition, for example, may consist of ambivalence, vacillation, or active suppression of one or the other dispositions.
Monotonicity	Strength of the response varies as a function of the extremity of the stimulus.	The functions, however, appear to be negatively accelerating rather than linear across the full dynamic range.
Antagonistic effects	Directional response effects of positive affect (approach) are generally opposite to those of negative effect (withdrawal).	Exceptions exist, such as when aversive threat is met with aggressive defense, or when a pain stimulus (e.g., flu shot) is approached and solicited.
Functional separability	Activation of positivity and negativity is partially separable.	This separation confers additional adaptability

		and flexibility for learned dispositions.
Modes of evaluate activation	Positivity and negativity can be activated reciprocally, uncoupled, or nonreciprocally.	At high levels of coactivation (which minimizes the dynamic range, reduces response lability, and maximizes directional flexibility), energy expenditure is taxing over long periods of time; eliciting circumstances tend to be avoided.
		The reciprocal activation postulate in prior models of affect and emotion is replaced by this postulate.
		Bivalent modes of evaluative action requires at least a three-dimensional approach: 1 for positivity; 1 for negativity; 1 representing the net behavioral predisposition or response orchestrated by the affect system.
Parallel evaluative processing	The ability to achieve coactivation of positivity and negativity by attending to positive and negative features of a stimulus simultaneously (e.g.,	Although there may be some reciprocal inhibition between the evaluative dimensions, they are at least partly independently expressed.

	bittersweet, disappointing wins).	
Low-pass filtering	The ability to achieve coactivation of positivity and negativity by oscillating between positive and negative stimuli with sufficient speed that results in the sustained activation of positivity and negativity.	Even though there can be an oscillation between positive and negative activation, if the speed of the presentation of the contrasting stimuli is faster than the low-pass filter cutoff, the activation of each cannot follow the speed of the oscillations and coactivation results (ambivalence)
		From both a practical and theoretical perspective, a rapidly switching oscillation between positive and negative (or more generally, from the two ends of a bipolar system) is equivalent to a bivariate system in its formal properties.
Distinct activation functions	The partial segregation of the positive and negative evaluative channels allows for distinctive activation functions for positivity and negativity.	The activation functions may be at least in part contextually dependent.
Positivity offset	The offset (intercept) for the positive activation function is higher than that of the negative activation function.	Motivation to approach is stronger than the motivation to withdraw at very low levels of evaluative

		activation; this promotes exploratory behavior – without a positivity offset, a person in a neutral environment is unlikely to approach novel stimuli.
Negativity bias	The gain for the negative activation function is higher than that of the positive activation function.	Motivation to withdraw is stronger than the motivation to approach at very high levels of evaluative activation; it is more difficult to overcome a fatal (or near fatal) assault than to return to an opportunity unpursued.
Recalibration	The activation functions for positivity and negativity are capable of the same kind of recalibrations based on salient contextual and accessible stimuli as is seen in receptor mechanisms.	Both sensitivity to small variations among stimuli and a dynamic range suitable to detect a wide array of affective stimuli are preserved.
Affective dispositions	There are measurable individual differences in the positivity offset and negativity bias.	These individual differences may have both a biological and an experiential/psychological basis.
Heterarchical organization	Evaluative processes are implemented at multiple, representative levels of the neuraxis.	There is a continuum of neuraxial organization that extends throughout the central nervous system in a heterarchical structure, ranging from the spinal cord to the frontal lobes.

Rostral, in contrast to caudal, neurobehavioral organizations are slower, more serial like; susceptible to more contextual control; potentiate greater response flexibility; and manifest multiple modes of appetitive and aversive activation. The multiple levels of processing can result in coordinated synergistic outcomes, or can lead to conflicts.

The ESM is based on the premise that the affect system has been shaped by the hammer and chisel of natural selection to produce a range of generally adaptive responses, and it posits that there are more conjunctions of factors that produce negativity than positivity. As Tolstoy observed in *Anna Karenina*, all happy families are alike one another, but each unhappy family is unhappy in its own way. When an event elicits positive affect or emotion, staying the course generally is sufficient. When negative affect or emotion is elicited, what constitutes an adaptive response may vary greatly across eliciting events. The ESM, therefore, posits that there is a greater diversity of adaptive responses for negative than positive antecedents, including a greater diversity of negative than positive emotions. Thus, the *heteroscedacity postulate* of the ESM states that the constellation of antecedents, emotions, expressions, and response is more diverse for negativity than for positivity (Cacioppo et al., 1997).

The ESM does not deny that valence is a central dimension in the affect system. A fundamental premise of the ESM is that the affect system evolved to help organisms differentiate hostile from hospitable stimuli and to respond adaptively to these stimuli. The ESM, however, posits that the valence dimension does not provide a sufficient depiction of the architecture of the affect system. We will return to additional postulates of the ESM in a subsequent section, but we begin by discussing what led to the ESM.

Personal Narrative of the Theory's Development

John Cacioppo and Richard Petty began working on attitudes in the early 1970s while they were graduate students together at Ohio State University. The field of attitudes at that time faced two challenges. First, independent variables appeared to have unreliable effects on attitude change; that is, variables like source credibility sometimes led to more attitude change, sometimes to less attitude change, and sometimes to no attitude change. Unspecified laboratory artifacts were thought to be a likely culprit. The implication was that carefully controlled laboratory research might have little or no bearing on the phenomenon theorists sought to understand – and that the voluminous body of attitude theory and research that had been spawned by careful experimental studies was of little scientific value. Second, attitudes were thought to not predict behavior. The implication was that attitudes and attitude change were not particularly important phenomena to understand because although attitudes might have a bearing on what people said, they putatively did not bear on what people actually did.

The Elaboration Likelihood Model (ELM) of attitudes and persuasion (Petty and Cacioppo, 1981, 1986) contributed to the resolution of these threats by specifying that attitude change was a multiply determined outcome; that is, there were different routes through which attitude development or change could be achieved, and that the same independent variable (e.g., source credibility) could produce various outcomes because its effects on attitude change and behavior depended on the specific route that was triggered by this variable in interaction with other variables (e.g., personal involvement, need for cognition). The ELM also specified the conditions under which these variables would operate to produce attitude change through the central or peripheral route. Moreover, the ELM predicted that attitudes [p. 55 ↓] were more likely to be temporally stable and guide behavior when these attitudes were formed through the central than peripheral route. Thus, attitudes predicted behavior but only under certain specifiable conditions.

Two decades later, Petty had expanded his interests to address issues such as attitude strength, attitude confidence, attitude anchoring effects, and the metacognitive

processes involved in attitudes and persuasion. Cacioppo, on the other hand, had become more interested in affect and stress, including the influence of affect and emotions on implicit social influence processes (e.g., emotional contagion; Hatfield et al., 1994) and the effect of somatovisceral processes on affect and attitudes (e.g., Cacioppo et al., 1993). During this period, Cacioppo returned to the psychology department at the Ohio State University as a faculty member in the social psychology program and in what is now called the psychobiology and behavioral neuroscience program. The director of the latter program was Gary Berntson, who similarly was interested in the topics of affect and stress. One of the first discussions between Cacioppo and Berntson was a paper the latter was writing on the inadequacy of the notion that the sympathetic and parasympathetic branches of the autonomic nervous systems were wired to be reciprocally activated. This insight led to discussions and collaborations that continue to this day, as well as to a theory of the structure and function of the autonomic nervous system (Berntson et al., 1991, 1993b) and to a companion theory of the structure and function of the affect system, the ESM (e.g., Berntson and Cacioppo, 2008; Cacioppo and Berntson, 1994, 1999; Cacioppo et al., 1997, 1999, 2004).

Theoretical developments in the ESM have been further fueled over the years by remarkable students and colleagues. Each has made many different contributions but it may be informative to mention a primary focus of each. Steve Crites performed a series of studies showing that evaluative and nonevaluative processing differed in the patterns of associated brain activation (Crites and Cacioppo, 1996; Cacioppo et al., 1996). Wendi Gardner (Cacioppo et al., 1997; Gardner and Cacioppo, 1995) and Jon Krosnick (Holbrook et al., 2001) provided early evidence for the functional separability of positive and negative affect. Tiffany Ito showed that the activation functions for positive and negative affect differed even when evaluative processing was spontaneous and unintended (Ito and Cacioppo, 2000; Ito et al., 1998b). Kyle Smith provided evidence that differences in the activation functions for positive and negative affect emerged early in the information processing stream to manifest as an attentional bias (Smith et al., 2003, 2006). Jeff Larsen has shown in various studies that positive and negative affect are not invariably reciprocally activated and helped articulate the conditions under which reciprocal activation would or would not occur (Larsen et al., 2001, 2004, 2009). Beth Crawford demonstrated that differences in the activation

functions of positive and negative affect have implicit effects on behavior, specifically that they influence the implicit learning of where appetitive and aversive stimuli are likely to be within a spatial context (Crawford and Cacioppo, 2002). Catherine Norris established that the nomothetic descriptions of the activation functions for positive and negative affect are robust, and that there are stable individual differences in the operating characteristics of the affect system – individual differences which predict subtle aspects of behavior (Norris and Cacioppo, 2009; Norris et al., in press). And Jackie Gollan has led the investigation of the possible clinical implications – specifically for understanding and treating affective disorders – of stable individual differences in the operating characteristics of the affect system (Gollan et al., 2009).

Work on the ESM has been funded primarily by the National Institute of Mental Health through a multi-institutional grant to the Center for the Study of Emotion and Attention led by Peter Lang at the University of Florida. Peter, Margaret Bradley at the University of Florida, Mike Davis at Emory [p. 56 ↓] University, Chris Patrick at the University of Minnesota, Arne Ohman at the Karolinska Institute, and the various affiliated investigators associated with this center have contributed to theoretical and empirical developments by being such constructive critics and colleagues over the years (e.g., Ito et al., 1998a; Schupp et al., 2000).

Intellectual History of the Theory

People have speculated about human affect and emotions for millennia, but the topic largely eluded scientific scrutiny until Louis Thurstone (Thurstone, 1928, 1931; Thurstone and Chave, 1929) adapted his quantitative methods in psychophysics to address the problem of measuring affect:

It has been stated by economists and by other social scientists that affect cannot be measured, and some of the fundamental theory of social science has been written with this explicit reservation. Our studies have shown that affect can be measured. In extending the methods of psycho-physics to the measurement of affect we see the possibility of a wide field of application by which it will be possible to apply the methods of quantitative scientific thinking to the study

of feeling and emotion, to aesthetics, and to social phenomena (Thurstone, 1931: 269).

Thurstone's methods included paired comparisons and rating scales, and his pioneering work led to the rating scales used in many surveys and experimental studies of affect today.

According to Thurstone's model, positive and negative affect fall on the opposite ends of a bipolar valence continuum, and this valence dimension is the central theoretical structure of the affect system:

The affect about an object may be of strong intensity or it may be weak. The positive and negative affect therefore constitutes a linear continuum with a neutral point or zone and two opposite directions, one positive and the other negative. Measurement along this affective continuum is of a discriminatory character with the discriminant error as a unit of measurement (Thurstone, 1931: 261).

More formally, affect in Thurstone's model is specified to be a joint function of positively and negatively valent activation functions (*evaluative activation postulate*); the strength of the response varies as a function of the extremity of the stimulus (*monotonicity postulate*); the responses to an equally spaced series of valent stimuli produces a linear output (*linearity postulate*); the directional response effects of positive affect (e.g., approach/withdrawal) is generally opposite to that of negative affect (*antagonistic effects postulate*); positively and negatively valent activation functions are reciprocally controlled (*reciprocal activation postulate*); and an increase in positive affect and a decrease in negative affect elicited by a stimulus produce the same movement and positioning along the valence continuum and therefore are indistinguishable in terms of the affect system (*functional equivalence postulate*). The valence structure in Thurstone's model could be viewed as operating analogously to a balance knob on a stereo to produce movements along this continuum ranging from "very positive and not at all negative" at one end to "very negative and not at all positive" at the other. Because the simple valence theory casts positive and negative affect as functionally equivalent – that is, the activation of positive affect is equivalent to the reduction in the activation of negative affect and vice versa – the theory further predicts that the activation function

for positive affect is the same as the activation function for negative affect (*equivalent activation functions postulate*).

Twenty years later, Neal Miller (1951, 1961) proposed that approach and withdrawal were behavioral manifestations with generally antagonistic effects on behavior that could come from *distinguishable* motivational substrates. His behavioral theory was enriched by conceptualizing approach and withdrawal separately, investigating their [p. 57 ↓] unique antecedents and consequences, and examining the behavioral constraints that led to approach and withdrawal tendencies. The division of motivational substrates into appetitive/nurturant/approach and aversive/defensive/withdrawal subsets and the view that appetition and aversion generally have antagonistic effects are central tenets of several contemporary theories of emotion, including Gray's (1987), Lang et al. (1990), Kahneman and Tversky's (1979), and Carver's (2001), as well as the ESM. All except the ESM, however, assumed or explicitly posited that appetition/approach and aversion/withdrawal were invariably reciprocally activated.

A theoretical model of the architecture and operating characteristics of the affect system makes it possible to generate new and falsifiable predictions, and this applies to the simple valence model of the affect system as well. For instance, it follows from this valence model that a neutral stimulus, one that arouses neither positive nor negative affect, is positioned in the middle of the valence continuum, but so is a stimulus that arouses strong positive and negative affect in equal measure – that is, an ambivalent stimulus. If the valence continuum were a sufficient description of the structure of the affect system, people would have difficulty differentiating neutral and ambivalent stimuli. Empirical evidence is inconsistent with this prediction (e.g., Edwards and Ostrom, 1971; Grabenhorst et al., 2007), which raises the possibility that a single valence continuum does not adequately capture the structure and operating characteristics of the affect system (cf. Cacioppo and Berntson, 1994).

Physical constraints generally restrict *behavioral* manifestations to bipolar actions (approach/withdrawal), but the ESM posits that evolution favors the animal that can learn, represent, and access rapidly whether approach or withdrawal is adaptive when confronted by a stimulus. There is a behavioral efficiency, a conservation of limited cognitive resources, and a reduction in physiological stress that is served by mental representations of general and enduring net action predispositions toward classes of

stimuli. Accordingly, the ESM posits that guides for one's actions in future encounters with the target stimuli, such as those provided by affective or emotional responses to the stimuli, will tend to be more expected and stable when organized in terms of a bipolar evaluative dimension (*energetic efficiency postulate*). The bipolar (positive/negative) structure may represent a stable endpoint, but it is not sufficient to model the structures or operations that preceded this adaptive endpoint.

Evidence that a person's mixed affective reactions migrate toward a bipolar organization dates back more than half a century ago. Brehm (1956) found that individuals, following a selection between two alternatives, spread the appeal of these alternatives by some combination of amplifying the positive features of the chosen alternative, diminishing the negative features of the chosen alternative, magnifying the negative features of the unchosen alternative, and minimizing the positive features of an unchosen alternative. This motivational push toward affective bipolarity was especially strong when subjects initially regarded the alternatives to be similarly appealing.

The fact that positive and negative affects have antagonistic effects on approach or withdrawal behavior does not mean that the affective inputs are invariably reciprocally activated. If you press your palms together with your forearms parallel to the floor, the force you exert with your right and left arms has antagonistic effects on the movement of your palms, but as you increase the force you exert with your right and left arms, your palms will not move as long as the forces you exert are equal. In this example, you have coactivated the muscles in your right and left arms rather than reciprocally activated them. The effects of activating the muscles of your right and left arms in this position are still antagonistic – pushing with your right arm moves your palms to the left and pushing with your left arm moves your palms to the right. But these antagonistic effects do not [p. 58 ↓] mandate that you necessarily exert reciprocal activation – increasing the force with which you are pushing with one arm while simultaneously decreasing the force with which you are pushing with the other.

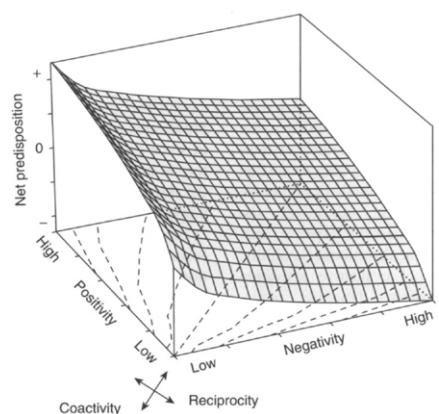
As we noted in the opening of the chapter, the ESM posits two abstract dimensions, positivity and negativity, each of which can range from zero to high levels of activation as a result of the activational status of its constituent elements. In addition, the ESM posits that affect is a joint function of positively and negatively valent activation functions (*evaluative activation postulate*); the strength of the response varies as a function of the

extremity of the stimulus (*monotonicity postulate*); the directional response effects of positive affect (e.g., approach/withdrawal) are generally opposite to that negative affect (*antagonistic effects postulate*); the activation of positivity and the activation of negativity are partially separable (*functional separability postulate*), and positivity; and negativity can be activated reciprocally (e.g., mutually exclusive, incompatible), uncoupled (e.g., singularly activated), or nonreciprocally (e.g., coactivational or coinhibitory) (*modes of evaluative activation postulate*). Thus, the ESM expands the principle of reciprocal evaluative activation to accommodate: (1) the separable activation of positivity and negativity, (2) the investigation of their unique antecedents and consequences, and (3) the concept and formal properties of modes of evaluative activation. Accordingly, the reciprocal activation postulate in prior models of affect and emotion is replaced by the modes of evaluative activation postulate.

The introduction of bivalent modes of evaluative activation requires at least a three-dimensional representation, one each to represent activation functions for positivity and negativity and a third “valence” dimension that represents the net behavioral predisposition or response orchestrated by the affect system. A two-dimensional representation of the activation of positivity and negativity, depicted as the bottom plane in [Figure 3.1](#), is termed the evaluative space and may provide a more comprehensive formulation for depicting these operations. For instance, this bivariate evaluative space accommodates all possible combinations of positive and negative evaluative activation: (1) the reciprocal mode of evaluative activation is represented as one diagonal vector that ranges from maximal positivity/minimal negativity to maximal negativity/minimal positivity, (2) the nonreciprocal mode of evaluative activation is represented as the alternate diagonal that ranges from minimal positivity and negativity to maximal positivity and negativity, and (3) the uncoupled modes of evaluative activation are represented as vectors lying along the axes. The family of vectors parallel to those above represent the general categories, or modes, of evaluative activation expressed from varying starting points within the two-dimensional plane depicted in [Figure 3.1](#).

Figure 3.1 Illustrative bivariate evaluative space and its associated affective response surface. This surface represents the net predisposition of an individual toward (+) or away from (-) the target stimulus. This net predisposition is expressed in relative units and the axis dimensions are in relative units of activation. The point on the surface overlying the left axis intersection represents a maximally positive predisposition, and

the point on the surface overlying the right axis intersection represents a maximally negative predisposition. Each of the points overlying the dashed diagonal extending from the back to the front axis intersections represent the same middling predisposition. Thus, the nonreciprocal diagonal on the evaluative plane – which represents different evaluative processes (e.g., neutral to ambivalence) – yields the same middling expression on the affective response surface. Dashed lines (including the coactivity diagonal) represent isocontours on the evaluative plane, which depict many-to-one mappings between the affective response surface and the underlying evaluative space. These isocontours are illustrative rather than exhaustive. (Adapted from Cacioppo and Berntson, 1994.)



Given the affect system evolved to guide behavior, information processed by the affect system does not stop with its registration on the evaluative space. Instead, the antagonistic effects of the activation of positivity and negativity are integrated into a net affective predisposition or action, which can be represented as an overlying bipolar response surface (see [Figure 3.1](#)). The positivity \times negativity evaluative space depicted in [Figure 3.1](#) represents the level of activation of the underlying positive and negative inputs to a bivalent affective response. The resulting behavioral predisposition to approach or withdraw can be represented in terms of an overlying surface whose projection on the z-axis constitutes the bipolar valence dimension. It is possible to derive this overlying surface for all combinations of positivity and negativity. Note that the mapping from the evaluative (positivity \times negativity) space to the valence dimension is many-to-one. The dashed lines in the evaluative space are illustrative, as the points

constituting any of the dashed lines (i.e., *iso-affective contours*) map into the same point on the bipolar valence dimension. Thus, knowledge of where one is in the evaluative space permits mapping to [p. 59 ↓] the valence dimension, but the reverse is not the case – except at the endpoints of the valence continuum (not at all positive and very negative; not at all negative and very positive). Therefore, the range of indeterminism is smallest at the anchors of the reciprocal diagonal (“very negative/not at all positive” and “very positive/not at all negative”) and largest along the coactivation diagonal.

Evolution can genetically endow only limited fixed adaptive responses relative to the potential range of circumstances an organism could encounter. Therefore, there is an evolutionary pressure to maximize flexibility and learning. According to the ESM, the partial segregation of the positive and negative evaluative channels in the affect system confers the additional flexibility of orchestrating appetitive and aversive motivational forces via modes of evaluative activation, which in turn affords greater flexibility for learned dispositions.

For instance, the concept of the “mode of evaluative activation” was derived from the architecture we posited for the affect system. [p. 60 ↓] These modes, in turn, produce formal response properties. The behavioral effects of positivity and negativity are generally antagonistic, so reciprocal activation produces a large dynamic range, high response lability, and high directional stability – the very qualities that are adaptive once a stimulus is determined to be hospitable or hostile. Contrary to the reciprocal mode of evaluative activation, coactivation minimizes the dynamic range, minimizes the response lability, and maximizes the directional flexibility. Uncoupled modes of evaluative activation produce intermediate dynamic range, response lability, and directional stability/flexibility. The energy expenditure in highly coactivated states is further posited to be taxing over long periods of time so, as noted above, coactivated states tend to be resolved or the eliciting circumstances tend to be avoided or denied. Consistent with this postulate, ambivalence has been associated with evaluative instability (e.g., Hass et al., 1991).

To illustrate, a thirsty animal on the Savannah would be motivated to go to the water hole for water (appetitive stimulus), where predators also lurk (aversive stimulus). In such circumstances, coactivation may be more adaptive, at least momentarily, as the animal remains in a highly energetic (i.e., prepared) state while it stoops to drink

and scans for predators ready to lurch from behind a bush or from beneath the murky water. Coactivation of positivity and negativity permits the thirsty animal to approach and consume some of the needed water while also maximizing the speed with which it can respond (reciprocally activate withdrawal and inhibit approach) when the predator approaches.

The stochastic and functional independence of positive and negative affect have been demonstrated (e.g., see Berntson and Cacioppo, 2008; Cacioppo and Gardner, 1999; Cacioppo et al., 1997, 2004). Some bipolar theorists have argued that positive and negative should not be expected to exhibit strong negative correlations due, for instance, to random and nonrandom measurement error (e.g., Green et al., 1993). However, these explanations for the stochastic separability documented in some studies cannot explain why the same measures and procedures would produce a bipolar structure in some circumstances and a bivalent structure in others – as suggested by our story of the thirsty animal on the Savannah. Is there empirical evidence of this sort?

Larsen et al. (2001) measured people's feelings of happiness and sadness – two emotions at opposite ends of the bipolar circumplex – during a normal day and a day characterized by complex feelings and emotions. For instance, they found that individuals were more likely to report feeling both happy and sad immediately after watching the film *Life Is Beautiful* (study 1), moving out of their college dormitories (study 2), and graduating from college (study 3) than in more typical situations (e.g., a typical day on campus). In addition to including standard emotion items such as happy and sad, in study 3, Larsen et al. included a more intriguing emotion: bittersweet. The term "bittersweet" implies a commingling of positive and negative feelings and is therefore difficult to place in the circumplex and other bipolar frameworks. Yet its inclusion in the lexicon suggests that it can sometimes characterize individuals' feelings. Consistent with this possibility, graduates were not only more likely to report feeling both happy and sad than nongraduates, they also reported feeling more bittersweet than nongraduates. In sum, Larsen et al. replicated the typical finding that happiness and sadness are largely mutually exclusive in routine, steady-state conditions (i.e., a normal day), but demonstrated that these two seemingly opposite emotions can co-occur under specifiable circumstances.

Additional research has shown that mixed feelings can occur not only in situations as rich as those studied by Larsen et al. (2001), but in situations as simple as a game of chance. In a laboratory experiment with a gambling task, Larsen et al. (2001) presented [p. 61 ↓] participants with 50/50 chances to win one of two amounts (e.g., \$7 or \$11) or lose one of two amounts. On winning trials, participants were more likely to feel both good and bad about winning the smaller outcome (i.e., a *disappointing win*) than the larger. Even though winning felt good, missing an opportunity to win an even greater amount felt bad. Similarly, participants were more likely to report mixed feelings after they lost the smaller of two amounts (i.e., *relieving losses*) rather than the larger. Thus, losing felt bad, even as avoiding a larger loss felt good.

The ESM is concerned with coactivation of underlying mechanism more so than the resulting experience of mixed feelings. One means of achieving this coactivation of positivity and negativity is by attending to positive and negative features (*parallel evaluative processing postulate*). Though attention is limited (e.g., Kahneman, 1973), the extant evidence supports the notion that attention can be directed to at least two channels of information (de Gelder and Vroomen, 2000; Spelke et al., 1976).

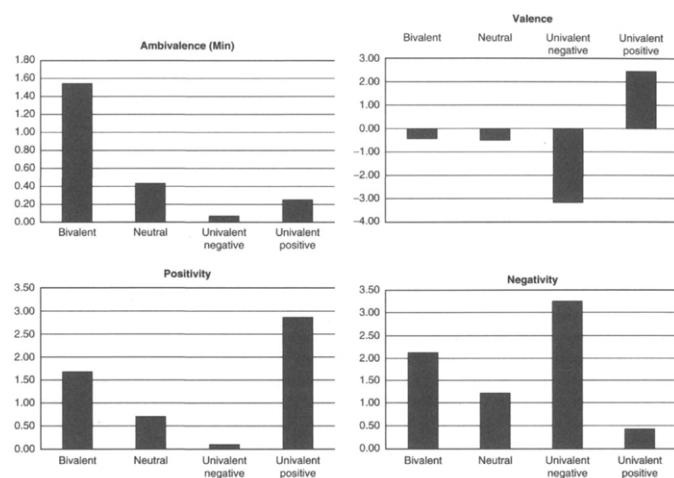
A second means of achieving coactivation posited by the ESM is oscillation between the positive and negative stimuli with sufficient speed that the low-pass filtering properties of the activation functions result in the sustained activation of positivity and negativity (*low-pass filtering postulate*). When one colloquially says that affect is sticky, the reference is to the low-pass filtering properties of the activation functions for positivity and negativity relative to cognition. A stimulus whose rate of oscillation exceeds a low-pass filter cutoff results in sustained activation. The theoretical implication is that the oscillation of positive and negative stimuli can produce a coactivation of positivity and negativity.

How the speed of the oscillation between two bipolar opposites can modulate activation and perception can be illustrated with a flip book, the pages of which alternate between solid black and solid white. If one stares for ten seconds at a black page and flips one page to the white page, the brightness is enhanced by a contrast effect. If one flips the pages slowly, the black and white pages are clearly visible. When the pages are flipped more quickly (e.g., 60 Hz), we no longer see the alternating black and white pages but instead we perceive the streaming pages to be gray. That is, even though there is an oscillation between high and low luminance, the speed of their presentation is too fast to

follow each presentation so we perceive a fusion of the two. There is a range of speeds for which this coactivation results. Too slow an oscillation, and the beginning and end of each stimulus is clearly perceived; too fast an oscillation, and nothing is perceived.

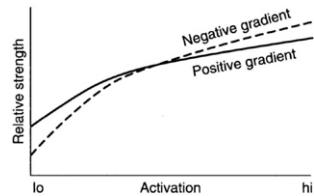
To determine whether the oscillation of univalent positive and univalent negative stimuli could produce evaluative coactivation, Cacioppo et al. (2009) presented brief, serial presentations of photographs of political figures that elicited univalent positive affect or univalent negative affect. Four conditions were constructed: (1) one of the oscillating pictures elicited univalent positive affect and the second elicited univalent negative affect (bivalent condition); (2) each of the two oscillating pictures elicited neutral affect (neutral condition); (3) each of the two oscillating pictures elicited negative affect (univalent negative condition); and (4) each of the two oscillating pictures elicited positive affect (univalent positive condition). The results, which are summarized in [Figure 3.2](#), confirmed that brief, serial presentations of photographs of political figures that elicited univalent positive affect or univalent negative affect led to the coactivation of positivity and negativity. These results suggest that one mechanism by which ambivalence can be aroused is oscillating attention between univalent positive and univalent negative features of stimuli ([Figure 3.2](#), upper left panel and bottom panels), and that the level of coactivation is not determinable from the valence dimension per se ([Figure 3.2](#), upper right panel).

Figure 3.2 Participants evaluated stimulus blocks consisting of pairs of targets they rated as positive, negative, or neutral. The alternating presentation of positive and negative stimuli (bivalent blocks) evoked higher levels of ambivalence than the neutral and univalent blocks. The bivalent blocks, however, show valence scores indistinguishable from those of the neutral blocks. This shows that conventional bipolar valence measures do not convey information about underlying positive and negative affect and may therefore mask ambivalent states. (From Cacioppo et al., 2009.)



As shown in [Figure 3.3](#), the ESM further posits that the partial segregation of the [p. 62 ↓] [p. 63 ↓] positive and negative evaluative channels in the affect system affords evolution of the opportunity to sculpt distinctive activation functions for positivity and negativity (*distinct activation functions postulate*), that both activation functions are negatively accelerating (*nonlinearity postulate*), that the intercept for the positive activation function (i.e., the approach motivation at zero input) is higher than the intercept for the negative activation function (*positivity offset postulate*), and that the gain for the negative activation function is higher than that for the positivity activation function (*negativity bias postulate*). The consequence of the positivity offset is that the motivation to approach is stronger than the motivation to withdraw at very low levels of evaluative activation, whereas the consequence of the negativity bias is that the motivation to withdraw is stronger than the motivation to approach at high levels of evaluative activation.

Figure 3.3 Activation functions for positive and negative dimensions of affective processing; the x-axis represents affective input, whereas the y-axis represents output of the system. The ESM proposes that there are two asymmetries in affective processing: the positivity offset is the result of greater positive than negative affect at low levels of emotional input; the negativity bias is the result of stronger responses to negative than to equally extreme positive input. (Adapted from Cacioppo and Berntson, 1994.)



The theoretical rationale for the positivity offset is that it produces exploratory behavior. Without a positivity offset, an organism in a neutral environment may be unmotivated to approach novel objects, stimuli, or contexts. The neophobic response to foreign stimuli that characterizes most species permits an initial period of observation. With no negative outcomes, this exposure allows the initial neophobic response to habituate, thereby allowing exploratory behavior to manifest. In the absence of such a motivation to explore, organisms would learn little about novel or neutral appearing environments and their potential reward value. With a positivity offset, however, an organism facing neutral or unfamiliar stimuli would be weakly motivated to approach, and with the quick habituation of the initial fear response, to engage in exploratory behavior. Such a pairing of initial neophobic and subsequent exploratory tendencies may have important survival value, at least at the level of a species. A positivity offset also fosters social cohesion even in the absence of other information about conspecifics.

The term “positivity bias” in the literature may be something of a misnomer. It has been used to refer to the finding that people about whom only neutral information is known are nevertheless rated positively. The term “bias” in engineering refers to the gain of an amplifier or activation function. We, therefore, reserved the use of the term “bias” to the “negativity bias,” which refers to differences in the gain of the activation functions for positivity and negativity, and we instead use the term “positivity offset” to refer to differences in the thresholds (i.e., intercepts) for these activation functions. The research on the “positivity bias” in person perception, however, may reflect the operation of the positivity offset in the affect system. For instance, Cacioppo et al. (1997) investigated the robust “positivity bias” in impression formation to determine whether it was limited to diagnosticity (cf. Skowronski and Carlston, 1989) or if it reflects a more general positivity offset in the affect system. The results of a series of studies confirmed that the positivity offset in impression formation was not limited by the social desirability concerns of the participants, by the type or diagnosticity of the

neutral behaviors used, nor by the similarity between target and participant. Indeed, the positivity offset was observed not only in impressions of human targets, but with impressions of novel fish and insects as well. The positivity offset demonstrated in this work, then, could not have been a result of the neutral behaviors, implying the absence of negative attributes (e.g., "Sam is susceptible to the laws of gravity"). Neither could it have merely reflected the process of similarity leading to attraction, as Sears' "person positivity bias" (1983) would have predicted. Instead, the positivity offset appeared to be a more general operating characteristic of the affect system.

The theoretical rationale for the negativity bias is that it is more difficult to overcome a fatal (or even a near-fatal) assault than to return to an opportunity unpursued, so it is more adaptive to err on the side of caution as threats get nearer. The negativity bias, therefore, provides an adaptive response function [p. 64 ↓] that complements the positivity offset. Human taste buds respond to sweet, salty, sour, and bitter stimuli. Most can detect sweetness in approximately one part in 200, saltiness in one part in 400, sourness in one in 130,000, and bitterness in one in 2,000,000. From the perspective of the affect system, a given amount of a negative or threat-related gustatory stimulus (e.g., most poisons taste bitter) activates a stronger affective response than the same amount of a positive (e.g., sweet) gustatory stimulus. This may be more than an epicurean curiosity; it may represent differences in the activation functions for positive and negative affective processing (see reviews by Baumeister et al., 2001; Cacioppo and Gardner, 1999; and Cacioppo et al., 1997). Moreover, the combination of spatial and affective information is essential for many approach and avoidance behaviors, and thus for survival. As predicted by the ESM, Crawford and Cacioppo (2002) found that the incidental learning of the likely spatial location of affective stimuli is greater for negative than positive stimuli.

According to the ESM, these distinct activation functions have evolved because they produce a neurobehavioral organization that is generally adaptive. There are two additional implications of this formulation that should be noted here. First, the activation functions for positivity and for negativity must balance the trade-off between dynamic range and sensitivity (precision). To illustrate, let us specify the dynamic range of the activation function for negativity to be an arbitrary and small amount, x . A weak negative stimulus (e.g., a predator at a safe distance) would produce a weak activation of negativity, but if the distance to the predator were to be reduced, the increase in

negativity would soon reach the maximum activation for negativity. This might produce very precise changes in activation with slight changes in distance to the predator (sensitivity), but it would also mean a predator that was nearby or one some distance away would produce maximal and comparable activation. This would mean the animal would avoid predators even when at a safe distance, dramatically reducing their access to food and water. Another alternative might be for the dynamic range of the activation function to be a very large amount, say $x \times 10^6$. While this would produce an activation function that would respond to changes in the distance from predators over a very large range, it would also decrease the sensitivity of these changes in that changes in negativity would require large rather than small changes in the distance to a predator. While such a range might permit distinctive activations to all or nearly all the negative stimuli an animal might encounter, this comes at a cost. Such a design could be fatal for an animal faced with multiple dangers in the same setting or when the animal's survival depends on its sensitivity to the proximity of a predator.

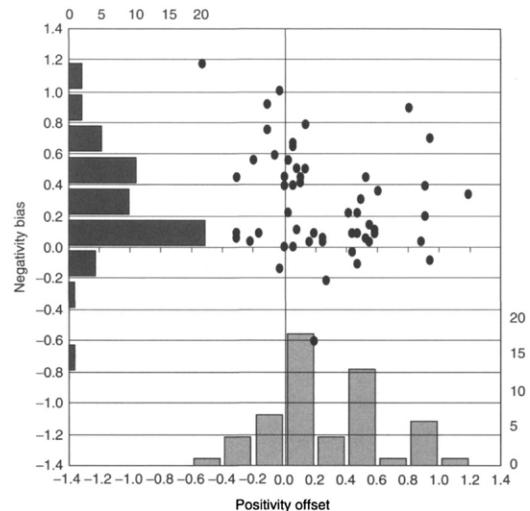
One of the means by which nature has solved this problem is adaptation/recalibration. If you have been sitting in a dimly lit room, you are able to see reasonably well once you have adapted. It is as if the dynamic range for luminance has been reduced so that you can see small variations in luminance that previously had been difficult to detect before your eyes adjusted (i.e., before the dynamic range for luminance was reduced). If someone suddenly turns on bright lights, the higher levels of luminance of the objects around you all produce comparably maximal levels of activation, making everything seem like an undifferentiated bright light, until your eyes again adapt to increase the dynamic range for luminance. The ESM posits that the activation functions for positivity and negativity are capable of the same kind of recalibrations based on the salient contextual and accessible stimuli (*recalibration postulate*). As a result of the recalibration of these activation functions, both sensitivity to small variations among stimuli and a dynamic range suitable to detect a wide array of affective stimuli are preserved, and so too is the positivity offset and the negativity bias.

Second, given that individual variation is the engine of natural selection, there should [p. 65 ↓] be measurable individual differences in the positivity offset and negativity bias (*affective dispositions postulate*). The underlying structure and operation of the affect system is generally outside people's awareness, and these dispositional tendencies are

similarly conceived as generally lying outside awareness, but like the affect system itself these dispositional tendencies should be measurable through people's responses to affective stimuli.

Consistent with this reasoning, stable and predictive individual differences in the positivity offset and the negativity bias have been identified (Ito and Cacioppo, 2005; Norris et al., *in press*). Participants in the Norris et al. study were exposed to three different sets of stimuli (pictures, sounds, and words), and during each set they were exposed to 66 stimuli, 6 of which were neutral and low in arousal, and 30 each of which vary in their extremity of pleasant or unpleasant and arousal but which were matched on these two dimensions. Ratings of each were made using the affect matrix – a 5 (positivity: zero to maximum) by 5 (negativity: zero to maximum) matrix on which participants rate each stimulus (Larsen et al., 2009). The positivity offset was indexed by the difference between the positivity and negativity ratings of the six neutral stimuli, and the negativity bias was gauged as the difference in rating of the six most extreme unpleasant stimuli minus the rating of the six most extreme (and initially matched on extremity and arousal) pleasant stimuli. Results revealed that individual differences in the positivity offset and negativity bias were uncorrected, temporally stable, and generalizable across ratings of pictures, sounds, and words (see [Figure 3.4](#)). Furthermore, individual differences in the positivity offset predicted the spatial learning for positive stimuli, whereas individual differences in the negativity bias predicted the spatial learning for negative stimuli. In sum, although most individuals exhibit both a positivity offset and a negativity bias, stable individual differences in these constructs predict what we learn about the world.

*Figure 3.4 Histograms of the aggregate measures of the negativity bias and positivity offset and a scatterplot depicting their relationship. Most participants exhibited a negativity bias ($M = 0.26$, $SE = 0.04$) and a positivity offset ($M = 0.27$, $SE = 0.05$), $t(64) = 6.46$ and 5.89 , respectively, $p < 0.001$, and the negativity bias and positivity offset were uncorrelated, $r(64) = -0.18$, ns. (From Norris et al., *in press*.)*



Finally, whereas many theories of affect assume evaluative processing is performed by a singular, or perhaps dual, processing mechanism, the ESM posits that positivity and negativity each represent the cumulative activation of multiple processing mechanisms. The ESM posits that affective states and responses are mediated by a network of distributed, often recursively connected, interacting neural regions in the central nervous system (including spinal cord reflexes), with the different areas making specific, often task-modulated contributions. Specifically, the ESM posits a continuum of neuraxial organization relevant to evaluative processing – a continuum that extends throughout the central nervous system in a heterarchical structure ranging from the frontal lobes to the spinal cord (*heterarchical organization postulate*).

The nineteenth-century neurologist John Hughlings Jackson emphasized the hierarchical structure of the brain, and the re-representation of functions at multiple levels within this neural hierarchy (Jackson, 1958/1884). The notion was that information is processed at multiple levels of organization within the nervous system. Primitive protective responses to aversive stimuli are organized at the level of the spinal cord, as is apparent in flexor (pain) withdrawal reflexes that can be seen even after spinal transection. These primitive protective reactions are expanded and embellished at higher levels of the nervous system (see Berntson et al., 1993a). More elaborate defensive/protective behaviors are organized within the brainstem, and decerebrate organisms (with no cerebral hemispheres) can display organized

escape and aggressive responses to noxious stimuli. The evolutionary development of even higher neural systems, such as the limbic system and cerebral cortex, endowed organisms with a further-expanded behavioral and motivational repertoire, that can capitalize on experience-dependent associative knowledge, information-processing networks, and cognitive strategies that anticipate and prepare for or avoid aversive encounters.

[p. 66 ↓] Evolution not only endowed us with primitive, lower-level adaptive reactions, but it sculpted the awesome information processing capacities of the highest levels of the brain. Thus, neurobehavioral mechanisms are not localized to a single level of organization within the brain, but are represented at multiple levels of the nervous system. At progressively higher levels of organization, there is a general expansion in the range and relational complexity of contextual controls and in the breadth and flexibility of discriminative and adaptive responses (Berntson et al., 1993a). Although higher-level systems confer greater behavioral variability and adaptive flexibility, they do not replace lower neurobehavioral mechanisms.

Adaptive flexibility of higher-level systems has costs, given the finite information-processing capacity of neural circuits. Greater flexibility implies a less rigid relationship between inputs and outputs, a greater range of information that must be processed, and a slower serial-like mode of processing. Consequently, the evolutionary layering of higher processing levels onto lower substrates has adaptive advantage in that lower and more efficient processing levels may continue to be utilized, and may be sufficient in some circumstances. For example, pain withdrawal reflexes, mediated by inherent spinal circuits, can manifest in rapid protective responses to pain stimuli. However, ascending pain pathways also convey information to higher **[p. 67 ↓]** levels of the neuraxis that subserve integrative aspects of affective, cognitive, and behavioral reactions such as fear, anxiety, avoidance, and/or aggression. Reflex responses provide a rapid low-level response, but they are not immutable, as higher neurobehavioral processes can come to suppress or bypass pain withdrawal reflexes (e.g., self-injecting insulin or recovering a billfold from a fire). These organizational features are not unique to defensive/protective behaviors, but rather reflect general neuroarchitectural principles that characterize generally hierarchical neural systems. Consequently, the defensive system and its re-representative organization across

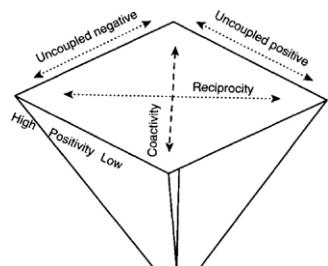
neuraxial levels offers a model system for conceptualizing neurobehavioral processes generally (e.g., see Berntson et al., 1998).

Although we have emphasized the features at the extremes, the heterarchical organization postulate of a continuum of neuraxial organization relevant to evaluative processing is in keeping with the principle of re-representation. Consider the architecture of spinal cord reflexes – the so-called final common pathway for behavior – where activation of flexor reflexes reciprocally inhibits extensor antagonists and vice versa. At the spinal cord per se, appetitive and aversive responses are reciprocally organized.

Recall that we defined the affect system as the efficient and manifold mechanisms that have evolved for differentiating hostile from hospitable stimuli and for organizing adaptive responses to these stimuli. The heterarchical nature of the central nervous system means that the functional architecture of the affect system extends downward from the bivariate evaluative plane at rostral levels to the reciprocal diagonal at the spinal cord wherein appetitive spinal cord reflexes and defensive spinal cord reflexes are reciprocally activated (see [Figure 3.5](#)). The ESM posits that affective states and responses are mediated by a network of distributed interacting neural regions in the central nervous system (including spinal cord reflexes). For instance, the spinal cord reflexes introduce a reciprocal bias in motor outputs, but this peripheral organization does not preclude the activation of both flexors and extensors (e.g., isometric contractions) via input from rostral brain areas (e.g., through voluntary efforts). Although the activities of the components of the affect system are generally integrated into a coherent cognitive/behavioral stream, the existence of multiple processing levels affords considerable flexibility in behavioral action as well as the potential for interference and conflict. In cases where relatively low-level processing is sufficient, [p. 68 ↓] or higher-level processing is precluded, lower substrates may predominate in behavioral expression. On the other hand, if higher cognitive computations, perhaps based on prior experience, raise the specter of a serious undesirable outcome, higher-level processes may predominate (e.g., recovering the billfold in the fire). In fact, although integrated to some extent, the multiple levels of processing may allow response conflicts, with different levels of processing each disposing toward differing behavioral responses. These conflicts may facilitate an outcome (e.g., retrieval of the billfold followed by a rapid, spinally mediated withdrawal from the fire), or may interfere (as in hesitancy,

vascillation, or indecision). In part, the latter arises from the fact that physical constraints preclude both actions, as the arm and hand cannot extend (reach out) and flex (pull back) at the same time. Physical constraints may belie the complexity of the underlying dispositions, however. Although the limb may not be able to extend and flex at the same time, the distinct underlying flexor and extensor muscles can, in fact, both contract at the same time. This may lead to inaction, but that inaction is not sufficient evidence for a lack of response dispositions. The limb response may be constrained along a single bipolar continuum of flexion or extension, but the underlying neuromuscular machinery is not so constrained and may reveal a broader fundamental bivariate structure to neurobehavioral control. Thus, the framework provided by the ESM has the potential to promote theory and research on the affect system both at the level of the individual components and at the level of the integrated network.

Figure 3.5 The bivariate evaluative space at the rostral levels extends downward to a bipolar line (the reciprocal diagonal) at the level of the spinal cord



The Theory's Applicability to Social Issues

The ESM has a variety of applications to social issues. Models of affect based on the valence continuum treat a reduction in negative affect as equivalent to an increase in positive affect. For instance, donor attitudes, intentions, and behaviors have typically been conceptualized as organized along a bipolar continuum. Research on blood, organ, and bone-marrow donor behavior suggests that negative beliefs and emotions may constitute a particularly difficult obstacle to inducing donor behaviors, and that increasing people's positive beliefs and emotions toward donor behavior is not sufficient to disinhibit donor behavior. The ESM, which does not treat positive and negative affect

as functionally equivalent, provides a framework and measurement methodology within which to study such phenomena.

The heterarchical structure proposed by the ESM also provides a framework for understanding implicit and explicit affective processes that may be playing a role in a variety of social problems, including racism. Briefly, latent inhibition in classical conditioning predicts that pre-exposure to a stimulus lessens the potency of classical conditioning, a prediction that has been confirmed in the human classical conditioning of affect (Cacioppo et al., 1992). If one assumes that children are exposed to majority more than to minority members, then salient distinguishing features of majority and minority members (e.g., skin coloration) of minority members should serve as a more powerful conditioning stimulus. Through the operation of latent inhibition, aversive depictions of whites and blacks in the news and media, even when equal, can lead to much stronger conditioned aversive responses to blacks than whites as a social category. These differential conditioned responses have no factual basis beyond the operation of latent inhibition, and the evocation of such feelings in the absence of a clear rationale promotes the search for an explanation and confabulation. When explicit rationalizations of this kind are challenged, the explicit beliefs can become less biased but this cognitive updating of facts and beliefs is not sufficient to fully eradicate conditioned emotional responses. That is, there are a multiplicity of evaluative representations within the affect system that reflect more than simple redundancy.

[p. 69 ↓] Finally, the ESM may offer unique contributions to our understanding of disorders of emotion such as depression, which, by its nature, is often profoundly interpersonal. Loss of reinforcement from the social environment, weakened social connections, reduced mastery of social situations, and uncontrollable social disruption are related to the onset of depression and its prediction of recurrence (Gortner et al., 1998; Lewinsohn et al., 1979). Moreover, social factors such as marital conflict, high levels of expressed negative emotion by family members, and low social support operate both as causes and consequences of depression (Hammen and Brennan, 2002). In this instance, initiating new affective responses to social cues, strengthening existing ones, or facilitating cognitive reappraisal to override attentional capture may modify depressive symptoms. However, it appears that these approaches and modification of social networks and behaviors alone are not sufficient to modify depression (Jacobson et al., 1996). The ESM may direct new integrative models that

identify individual differences in affective processes, which, in their own right, may identify social triggers of depression.

The independence of the positivity offset and negativity bias indices, as well as the separability of positive and negative affective processes outlined by the ESM, may identify the social processes associated with depression. Research indicates that although both depression and anxiety are characterized by increased negative affect, they can be distinguished by the addition of decreased *positive* affect in depression (Watson et al., 1988). For example, it is theoretically possible that individuals who exhibit greater relative negativity bias and lower positivity offset activity may experience more negative affect, respond to a negative mood induction with more negative affect, and report a greater propensity for behavioral/social inhibition. That is, behavioral inhibition interferes with efforts to cultivate socially reinforcing environments and networks, which may inadvertently increase the risk for depression onset.

Moreover, depressed individuals may be distinguishable from nonpsychiatric healthy control subjects in their relatively higher endorsement of negativity bias and lower endorsement of positivity offset. Conversely, through the lower activation of the negativity bias and higher activation of positivity offset, an individual may seek to generate novel and appealing social contexts that are sufficient to encourage mastery and pleasure. These individuals may experience more positive affect, respond with more positive affect to social stimuli, and have a greater propensity for environmental and social exploration. These initiatives may reap dividends in promoting social support and warding off depressive symptoms. Research has yet to determine whether the relative difference in positivity offset and negativity bias changes with symptom remission, but the ESM offers a theoretical blueprint for distinguishing affective processes within affective disorders.

Conclusion

Most readers will be familiar with Occam's Razor – that theoretical entities should not be multiplied unnecessarily. Perhaps not all readers will be familiar with Einstein's Razor, however, which states that theorists should make everything as simple as possible, but not simpler. Our goal in the ESM is to begin to delineate details of the functional

architecture and operating characteristics of the affect system. For instance, the ESM posits that the simple bipolar valence continuum, even when complemented by an arousal dimension as in the circumplex model, is not a comprehensive model of the affect system. The ESM instead posits that the outputs of the evaluative processors comprising the affect system determine bivalent action tendencies and actions. Such an organization fosters function: free and swift approach to appetitive stimuli and rapid and unfettered withdrawal from aversive stimuli.

[p. 70 ↓]

Acknowledgments

Preparation of this article was supported by grants from the National Institute of Mental Health (Grant No. P50 MH72850).

References

- Baumeister, R.F., Bratslavsky, E., Finkenauer, C. and Vohs, K.D. (2001) Bad is stronger than good . *Review of General Psychology* , 5(4), 323–370. <http://dx.doi.org/10.1037/1089-2680.5.4.323>
- Berntson, G.G., Boysen, S.T. and Cacioppo, J.T. (1993a) Neurobehavioral organization and the cardinal principle of evaluative bivalence . *Annals of the New York Academy of Sciences* , 702, 75–102. <http://dx.doi.org/10.1111/j.1749-6632.1993.tb17243.x>
- Berntson, G.G. and Cacioppo, J.T. (2008) The neuroevolution of motivation . In J. Shah, ed. and W. Gardner (eds), *Handbook of Motivation Science* , pp. 188–200. New York: Guilford.
- Berntson, G.G., Cacioppo, J.T. and Quigley, K.S. (1991) Autonomic determinism: The modes of autonomic control, the doctrine of autonomic space, and the laws of autonomic constraint . *Psychological Review* , 95(4), 459–487. <http://dx.doi.org/10.1037/0033-295X.98.4.459>

Berntson, G.G., Cacioppo, J.T. and Quigley, K.S. (1993b) Cardiac psychophysiology and autonomic space in humans: Empirical perspectives and conceptual implications . Psychological Bulletin , 114(2), 296–322. <http://dx.doi.org/10.1037/0033-2909.114.2.296>

Berntson, G.G., Sarter, M. and Cacioppo, J.T. (1998) Anxiety and cardiovascular reactivity: The basal forebrain cholinergic link . Behavioural Brain Research , 94(2), 225–248. <http://dx.doi.org/10.1016/S0166-4328%2898%2900041-2>

Brehm, J.W. (1956) Postdecision changes in the desirability of alternatives . The Journal of Abnormal and Social Psychology 52(3), 384–389. <http://dx.doi.org/10.1037/h0041006>

Cacioppo, J.T. and Berntson, G.G. (1994) Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates . Psychological Bulletin , 115(3), 401–423. <http://dx.doi.org/10.1037/0033-2909.115.3.401>

Cacioppo, J.T. and Berntson, G.G. (1999) The affect system: Architecture and operating characteristics . Current Directions in Psychological Science , 8(5), 133–137. <http://dx.doi.org/10.1111/1467-8721.00031>

Cacioppo, J.T., Crites, S. L, Jr. and Gardner, W.L (1996) Attitudes to the right: Evaluative processing is associated with lateralized late positive event-related brain potentials . Personality and Social Psychology Bulletin , 22(12), 1205–1219. <http://dx.doi.org/10.1177/01461672962212002>

Cacioppo, J.T. and Gardner, W.L. (1999) Emotion . Annual Review of Psychology 50, 191–214. <http://dx.doi.org/10.1146/annurev.psych.50.1.191>

Cacioppo, J.T., Gardner, W.L. and Berntson, G.G. (1997) Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space . Personality and Social Psychology Review , 1(1), 3–25. http://dx.doi.org/10.1207/s15327957pspr0101_2

Cacioppo, J.T., Gardner, W.L. and Berntson, G.G. (1999) The affect system has parallel and integrative processing components: Form follows function . *Journal of Personality and Social Psychology* 76(5), 839–855. <http://dx.doi.org/10.1037/0022-3514.76.5.839>

Cacioppo, J.T., Larsen, J.T., Smith, N.K. and Berntson, G.G. (2004) The affect system: What lurks below the surface of feelings? In A.S.R. Manstead, ed. , N.H. Frijda, ed. and A.H. Fischer (eds), *Feelings and Emotions: The Amsterdam Conference* , pp. 223–242. New York: Cambridge University Press.

Cacioppo, J.T., Marshall-Goodell, B.S., Tassinary, L.G. and Petty, R.E. (1992) Rudimentary determinants of attitudes: Classical conditioning is more effective when prior knowledge about the attitude stimulus is low than high . *Journal of Experimental Social Psychology* , 28(3), 207–233. <http://dx.doi.org/10.1016/0022-1031%2892%2990053-M>

Cacioppo, J.T., Priester, J.R. and Berntson, G.G. (1993) Rudimentary determinants of attitudes. II: Arm flexion and extension have differential effects on attitudes . *Journal of Personality and Social Psychology* , 65(1), 5–17. <http://dx.doi.org/10.1037/0022-3514.65.1.5>

Cacioppo, J.T., Rasinski, K.A., Sauerbrei, B.A. and Brehm, J.J. (2009) Ambivalence in the political arena . Unpublished manuscript.

Carver, C.S. (2001) Affect and the functional bases of behavior: On the dimensional structure of affective experience . *Personality and Social Psychology Review* , 5(4), 345–356. http://dx.doi.org/10.1207/S15327957PSPR0504_4

Crawford, L.E. and Cacioppo, J.T. (2002) Learning where to look for danger: Integrating affective and spatial information . *Personality and Social Psychology Review* , 13(5), 449–453.

Crites, S.L.J, and Cacioppo, J.T. (1996) Electrocortical differentiation of evaluative and nonevaluative categorizations . *Psychological Science* , 7(5), 318–321. <http://dx.doi.org/10.1111/j.1467-9280.1996.tb00381.x>

de Gelder, B. and Vroomen, J. (2000) The perception of emotions by ear and by eye . Cognition and Emotion , 14(3), 289–311. <http://dx.doi.org/10.1080/026999300378824>

Edwards, J.D. and Ostrom, T.M. (1971) Cognitive structure of neutral attitudes . Journal of Experimental Social Psychology , 7(1), 36–47. <http://dx.doi.org/10.1016/0022-1031%2871%2990053-9>

Gardner, W.L and Cacioppo, J.T. (1995) Multi-gallon blood donors: Why do they give? Transfusion , 35(10), 795–798. <http://dx.doi.org/10.1046/j.1537-2995.1995.351096026358.x>

Gollan, J.K., Norris, C., Heyne, M. and Cacioppo, J.T. (2009) Translating basic science to the clinical context: Improving outcome in behavioral treatment of depression . Unpublished manuscript.

Gortner, E.T., Gollan, J.K., Dobson, K.S. and Jacobson, N.S. (1998) Cognitive-behavioral treatment for depression: Relapse prevention . Journal of Consulting and Clinical Psychology 66(2), 377–384. <http://dx.doi.org/10.1037/0022-006X.66.2.377>

Grabenhorst, F., Rolls, E.T., Margot, C., da Silva, M.A.A.P. and Velazco, M.I. (2007) How pleasant and unpleasant stimuli combine in different brain regions: Odor mixtures . Journal of Neuroscience , 27(49), 13532–13540. <http://dx.doi.org/10.1523/JNEUROSCI.3337-07.2007>

Gray, J.A. (1987) The Psychology of Fear and Stress (2nd edition). Cambridge: Cambridge University Press.

Green, D.P., Goldman, S.L and Salovey, P. (1993) Measurement error masks bipolarity in affect ratings . Journal of Personality and Social Psychology 64(6), 1029–1041. <http://dx.doi.org/10.1037/0022-3514.64.6.1029>

Green, D.P., Salovey, P. and Truax, K.M. (1999) Static, dynamic, and causative bipolarity of affect . Journal of Personality and Social Psychology , 76(5), 856–867. <http://dx.doi.org/10.1037/0022-3514.76.5.856>

Hammen, C. and Brennan, P.A. (2002) Interpersonal dysfunction in depressed women: Impairments independent of depressive symptoms . *Journal of Affective Disorders* , 72(2), 145–156. <http://dx.doi.org/10.1016/S0165-0327%2801%2900455-4>

Hass, R.G., Katz, I., Rizzo, N., Bailey, J. and Eisenstadt, D. (1991) Cross-Racial Appraisal as Related to Attitude Ambivalence and Cognitive Complexity . *Personality and Social Psychology Bulletin* , 17, 83–92. <http://dx.doi.org/10.1177/0146167291171013>

Hatfield, E., Cacioppo, J.T., and Rapson, R.L. (1994). *Emotional Contagion* . New York: Cambridge University Press.

Holbrook, A.L., Krosnick, J.A., Visser, P.S., Gardner, W.L. and Cacioppo, J.T. (2001) Attributes toward presidential candidates and political parties: Initial optimism, inertial first impressions, and a focus on flaws . *American Journal of Political Science* , 45(4), 930. <http://dx.doi.org/10.2307/2669333>

Ito, T. and Cacioppo, J. (2005) Variations on a human universal: Individual differences in positivity offset and negativity bias . *Cognition and Emotion* , 19(1), 1–26. <http://dx.doi.org/10.1080/02699930441000120>

Ito, T.A. and Cacioppo, J.T. (2000) Electrophysiological evidence of implicit and explicit categorization processes . *Journal of Experimental Social Psychology* , 36(6), 660–676. <http://dx.doi.org/10.1006/jesp.2000.1430>

Ito, T.A., Cacioppo, J.T. and Lang, P.J. (1998a) Eliciting affect using the international affective picture system: Trajectories through evaluative space . *Personality and Social Psychology Bulletin* , 24(8), 855–879. <http://dx.doi.org/10.1177/0146167298248006>

Ito, T.A., Larsen, J.T., Smith, N.K. and Cacioppo, J.T. (1998b) Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations . *Journal of Personality and Social Psychology* , 75(4), 887–900. <http://dx.doi.org/10.1037/0022-3514.75.4.887>

Jackson, J.H. (1958/1884) Evolution and dissolution of the nervous system (Croonian Lectures) . In J. Taylor (ed.) Selected Writings of John Hughlings Jackson (Vol. 2) . New York: Basic Books.

Jacobson, N.S., Dobson, K.S., Truax, P.A., Addis, M.E., Koerner, K., Gollan, J.K. et al. (1996) A component analysis of cognitive-behavioral treatment for depression . *Journal of Consulting and Clinical Psychology* , 64(2), 295–304. <http://dx.doi.org/10.1037/0022-006X.64.2.295>

Kahneman, D. (1973) Attention and Effort . Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D. and Tversky, A. (1979) Prospect theory: An analysis of decision under risk . *Econometrica* , 47(2), 263–292. <http://dx.doi.org/10.2307/1914185>

Lang, P.J., Bradley, M.M. and Cuthbert, B.N. (1990) Emotion, attention, and the startle reflex . *Psychological Review* , 97(3), 377–395. <http://dx.doi.org/10.1037/0033-295X.97.3.377>

Larsen, J.T., McGraw, A.P. and Cacioppo, J.T. (2001) Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology* , 81(4), 684–696. <http://dx.doi.org/10.1037/0022-3514.81.4.684>

Larsen, J.T., McGraw, A.P., Meilors, B.A. and Cacioppo, J.T. (2004) Research article the agony of victory and thrill of defeat mixed emotional reactions to disappointing wins and relieving losses . *Psychological Science* , 15(5), 325–330. <http://dx.doi.org/10.1111/j.0956-7976.2004.00677.x>

Larsen, J.T., Norris, C.J., McGraw, A.P., Hawley, L.C. and Cacioppo, J.T. (2009) The evaluative space grid: A single-item measure of positivity and negativity . *Cognition and Emotion* , 23, 453–480. <http://dx.doi.org/10.1080/02699930801994054>

LeDoux, J.E. (2000) Emotion circuits in the brain . *Annual Review of Neuroscience* , 23(1), 155. <http://dx.doi.org/10.1146/annurev.neuro.23.1.155>

Lewinsohn, P.M., Youngren, M.A. and Grosscup, S.J. (1979) Reinforcement and depression . In R.A. Dupue (ed.), *The Psychobiology of Depressive Disorders: Implications for the Effects of Stress* , pp. 291–316. New York: Academic Press.

Miller, N.E. (1951) Learnable drives and rewards . In S.S. Stevens (ed.), *Handbook of Experimental Psychology* , pp. 435–472. New York: Wiley.

Miller, N.E. (1961) Analytical studies of drive and reward . *American Psychologist* , 16(12), 739–754. <http://dx.doi.org/10.1037/h0047048>

Norris, C.J. and Cacioppo, J.T. (2009) An fMRI investigation of neural mechanisms underlying the negativity bias . Unpublished manuscript.

Norris, C.J., Larsen, J.T., Crawford, L.E. and Cacioppo, J.T. (in preparation) Better (or worse) for some than others: Individual differences in the positivity offset and negativity bias . *Journal of Research in Personality* . Unpublished manuscript.

Petty, R.E. and Cacioppo, J.T. (1981) Issue involvement as a moderator of the effects on attitude of advertising content and context . *Advances in Consumer Research* , 8, 20–24.

Petty, R.E. and Cacioppo, J.T. (1986) The elaboration likelihood model of persuasion . *Advances in Experimental Social Psychology* , 19, 123–205. <http://dx.doi.org/10.1016/S0065-2601%2808%2960214-2>

Russell, J.A. (1980) A circumplex model of affect . *Journal of Personality and Social Psychology* , 39(6), 1161–1178. <http://dx.doi.org/10.1037/h0077714>

Schupp, H.J., Cuthbert, B.N., Bradley, M.M., Cacioppo, J.T., Ito, T. and Lang, P.J. (2000) Affective picture processing: The late positive potential is modulated by motivational relevance . *Psychophysiology* , 37(2), 257–261. <http://dx.doi.org/10.1111/1469-8986.3720257>

Sears, D.O. (1983) The person-positivity bias . *Journal of Personality and Social Psychology* , 44(2), 233–250. <http://dx.doi.org/10.1037/0022-3514.44.2.233>

Skowronski, J.J. and Carlston, D.E. (1989) Negativity and extremity biases in impression formation: A review of explanations . *Psychological Bulletin* , 105(1), 131–142. <http://dx.doi.org/10.1037/0033-2909.105.1.131>

Smith, N.K., Cacioppo, J.T., Larsen, J.T. and Chartrand, T.L. (2003) May I have your attention, please: Electrocortical responses to positive and negative stimuli . *Neuropsychologic* 41(2), 171–183. <http://dx.doi.org/10.1016/S0028-3932%2802%2900147-1>

Smith, N.K., Larsen, J.T., Chartrand, T.L., Cacioppo, J.T., Katafiasz, H.A. and Moran, K.E. (2006) Being bad isn't always good: Affective context moderates the attention bias toward negative information . *Journal of Personality and Social Psychology* , 90(2), 210–220. <http://dx.doi.org/10.1037/0022-3514.90.2.210>

Spelke, E., Hirst, W. and Neisser, U. (1976) Skills of divided attention . *Cognition* , 4(3), 215–230. <http://dx.doi.org/10.1016/0010-0277%2876%2990018-4>

Thurstone, L.L. (1928) Attitudes can be measured . *American Journal of Sociology* , 33(2), 529–554. <http://dx.doi.org/10.1086/214483>

Thurstone, L.L. (1931) The measurement of attitudes . *Journal of Abnormal Psychology* , 26, 249–269. <http://dx.doi.org/10.1037/h0070363>

Thurstone, L.L. and Chave, E.L. (1929) *The Measurement of Attitude* . Chicago: University of Chicago Press.

Watson, D., Clark, L.A. and Tellegen, A. (1988) Development and validation of brief measures of positive and negative affect: The PANAS scales . *Journal of Personality and Social Psychology* , 54(6), 1063–1070. <http://dx.doi.org/10.1037/0022-3514.54.6.1063>

Watson, D. and Tellegen, A. (1985) Toward a consensual structure of mood . *Psychological Bulletin* , 98(2), 219–235. <http://dx.doi.org/10.1037/0033-2909.98.2.219>

[p. 73 ↓]

<http://dx.doi.org/10.4135/9781446249215.n4>

Handbook of Theories of Social Psychology: Volume 1

Accessibility Theory

Contributors: E. Tory Higgins

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Accessibility Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n5>

Print pages: 75-97

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n5>

[p. 75 ↓]

Chapter 4: Accessibility Theory

E. Tory Higgins, ed.

Abstract

Over 30 years have passed since I began studying the priming and accessibility of social constructs. What has made the journey since then so enjoyable are all the colleagues and friends who collaborated in deepening and broadening our knowledge of the nature and functions of priming and accessibility. Writing this chapter helped me to appreciate how much we have learned over the years about the sources and the consequences of accessibility as a basic principle of psychology. We know, for example, that when recent and frequent priming are combined with chronic accessibility (the additivity principle), even extremely vague information about someone will be used as a basis for forming an impression of that person (the compensation principle). We know that when people are aware of the priming event, they will often correct for its influence, thereby producing an opposite judgmental bias (contrast). We know that subliminally priming social category knowledge will produce actions in line with that knowledge (even in the absence of a member of the category), but the specific action will also depend on someone's attitude toward category members. This latter lesson is part of a recent and developing story that priming and accessibility effects depend on the *relevance* of a stimulated construct. Priming and accessibility is not just about cognition. It is about motivation.

Accessibility Theory

When I was a social psychology graduate student at Columbia in the early 1970s, most of the seminars I took were taught by Stanley Schachter, but I ended up doing most of my research in collaboration with Janellen Huttenlocher who was then a professor at Columbia Teachers College. Together we did research in psycholinguistics; more specifically, research on verbal reasoning. My dissertation advisor was Robert Krauss and my dissertation was on developmental and social class differences in communication. Given all this, when I applied for a job at Princeton, the faculty members were not quite sure what I was. Because of my background with Schachter and Krauss, I was partly a social psychologist, but my overarching research issue – the relations among language, thought, and society – was not what interested most social psychologists. Because of my research on psycholinguistics and communication, I was partly a cognitive psychologist, but my formal graduate training was in social psychology.

[p. 76 ↓] So what did Princeton do? Princeton's Solomon-like decision was to hire me in their "open" slot (rather than in either social psychology or cognitive psychology) where I would teach developmental psychology courses, without my ever having taken even one course in developmental psychology. The rationale was simple, however. After all, wasn't Janellen Huttenlocher herself a developmental psychologist (even though our research together was *not* developmental)? And wasn't my dissertation on developmental differences in communication? Well, I guess they had me there!

I relate this story because I believe that it is precisely my background as a part social psychologist and part cognitive psychologist that was critical to my early research on priming and accessibility. Indeed, before going to Columbia for my PhD I had worked on a social psychology MA at the London School of Economics and Political Science. "Social psychology MA" you say? Aha! That means that you *were* a social psychologist! Actually, it is not that simple. My advisor at LSE, Norman Hotopf, was a PhD student of Bertrand Russell and a postdoctoral student with Frederick Bartlett. He is best known for his work on eye movement and on the relation between language and thought. His hero was Roger Brown. Once again, he was not a typical social psychologist and was, if anything, more of a psycholinguist. But before I left LSE in 1968 he had said to me

that what I should do in the future is to try to combine social psychology and cognitive psychology: "You know, something like *social cognition*."

Well, it took several years before I followed his advice, but once I reached Princeton in 1972, I decided to try to do exactly what he had told me – to combine social psychology and cognitive psychology. I had been interested in person perception for some time and was familiar with the "New Look" perspective on perception. The New Look literature had reported how individuals' needs and expectancies could influence their judgments of things, such as a greater need for money being associated with judging coins to be bigger in size. At the general level, what this work seemed to demonstrate ("seemed" because the conclusions were still somewhat controversial) was that individual difference factors could influence judgments separately from the properties of the target of judgment. In Burner's (1957a, b) terms, individuals can be set or *ready* to perceive certain things because they want them or expect them to be there, and this readiness results in their *going beyond the information given* in their judgments; that is, they make a judgment of the target that is not based solely on the target's properties.

According to Bruner (1957a: 133), categories varied in their accessibility, which is "the readiness with which a stimulus input with given properties will be coded or identified in terms of a category." To use Burner's (1957a: 130) example, if the category "apple" has high accessibility, then "apples will be more easily and swiftly recognized, a wider range of things will be identified or misidentified as apples, and in consequence the correct or best fitting identity of these other inputs will be masked." What this means is that the likelihood that some input will be categorized in terms of a given category depends not only on the overlap between the sensory input and the category's specifications but also on the accessibility of that category.

Where does accessibility come from? This, of course, was a key question. According to Bruner (1957a), the relative accessibility of a category depends upon two factors. First, accessibility depends on the *expectancies* of a person regarding the likelihood of a *type of event that will be encountered* in the environment, with higher accessibility for events of higher likelihood. High accessibility from such frequency of past exposure functions to *minimize the surprise value of the environment*. Second, accessibility depends on the *search requirements* of the person imposed by that person's *needs and ongoing*

enterprises. High accessibility in this case functions to *maximize the attainment of sought-after objects and events.*

[p. 77 ↓] It occurred to me that there was something about Burner's conceptualization of accessibility that was incomplete. According to this theory, a category was accessible in the present when we expected or wanted an instance of that category to appear. When a category was accessible, we looked for and searched for an instance of that category in order to avoid surprises or attain desired end-states. I was not satisfied with this account of accessibility. I was not satisfied because it did not cover the case of accessibility which Janellen Huttenlocher and I had uncovered a few years earlier.

After arriving at Columbia in 1968, I began working with Huttenlocher on explaining why different kinds of three-term syllogisms varied in difficulty – the psychology of verbal reasoning. The so-called “three-term series problem” involves two premises and a question, such as: “Tom is lighter than Bob”; “Dick is heavier than Bob,” “Who is heaviest?” The difficulty of solving the problems depends on the linguistic form of each premise (e.g., contains a marked adjective [light] versus an unmarked adjective [heavy]; is a regular form versus a negative equative (“Bob is not as heavy as Dick”)). What Janellen and I most cared about was to explain how the information from the two premises was combined, and how the linguistic form of each premise influenced the process of combining the information. But we got involved in a debate with Herb Clark about verbal reasoning that forced us to provide some explanation for the effect of the adjective in the question – “Who is heaviest?” versus “Who is lightest?” Janellen and I found that participants solved the problem faster when the adjective in the question matched the active in the *second* premise, that is, the most recent premise. We proposed that activation of the category meaning of the adjective in the second premise made that meaning temporarily more accessible, which in turn made it easier to activate that meaning again when the same adjective appeared in the question (Huttenlocher and Higgins, 1971) – what we would now call a *recent priming* effect on *temporary accessibility*.

In our studies, which involved many trials for each participant, the likelihood that the adjective in the question would be a mismatch to the adjective in the second premise was the same as the likelihood that it would be a match. Given this equal likelihood, having higher accessibility of the adjective in the second premise would not function

to “minimize the surprise value of the environment.” In addition, the goal of the task was to search for the *person* who was the answer to the question (e.g., “Tom”) and not to satisfy some need associated with the adjective in the second premise. That is, comprehending the question faster because its adjective matched the adjective in the second premise was not due to the question adjective being a desired end-state. What was clear to me was that something about accessibility was going on in the three-term series problem which was not captured in Bruner’s account. But what was it?

Much of the accessibility studied by Bruner and others in the New Look concerned relatively long-term differences in accessibility, such as differences in the value of a coin as a function of personal wealth. Some of the phenomena examined, such as food deprivation effects on perceiving food-related items, were more temporary, but even here it concerned hours of deprivation and not seconds. In contrast, the effect of the second premise adjective on comprehending the question adjective would be an accessibility effect after seconds of activation. Indeed, we knew that the effect of adjective match was reduced if participants had to say three digits between the second premise and the question (see Huttenlocher and Higgins, 1971). This is why we called it *temporary accessibility*. How might such *temporary accessibility* be conceptualized?

Huttenlocher and I had suggested that activating categorical meaning from exposure to a word (i.e., priming) made that meaning temporarily more accessible, which in turn made that meaning easier to activate upon later exposure to a meaning-related stimulus. In our case, the later meaning-related stimulus [p. 78 ↓] was the same adjective appearing again. But what if something more general were going on? What if recent activation of category meaning from priming a word made that category meaning temporarily more accessible, which in turn made it easier to activate upon later exposure to *behavioral information* that related to that meaning? If so, then verbal priming of a category in one situation could increase the likelihood of that category being used subsequently to categorize someone’s behavior in a separate situation – without any awareness of the first situation having influenced the subsequent social judgment. And what if the target person’s behavior was evaluatively ambiguous; that is, had features that fit two different categories with opposite valence (e.g., a persistent person versus a stubborn person)? If so, then perceivers could end up with either a positive or negative *impression* of the target person depending on whether – in an earlier totally separate situation – they had been verbally primed with one word or

another (i.e., the word “persistent” versus the word “stubborn”). This would be a very different kind of accessibility effect!

My First Priming and Accessibility Study: The “Donald” Study

Psychologists had long recognized that categorizing an object or person in a certain way had subsequent effects on how that person would later be remembered and evaluated (e.g., Carmichael et al., 1932; Kelley, 1950). But where did such categorization come from? The standard answer was some combination of the subjective perception of the target's properties (i.e., the perceived overlap between those perceived target properties and the properties of different stored categories), plus some motivational biases or preferences for applying one or another category to a particular target. Temporary accessibility of a category meaning from recent verbal priming was a *very different* kind of source. This source had nothing to do with the target's properties or the perceivers' needs. It concerned simply prior activation of the category meaning, which could occur in an *incidental* situation that had nothing to do with the target *and* could impact how the target is categorized without the perceiver being aware of the influence. It was the kind of scary source that reminded me of Freud's unconscious motives as a biasing factor in judgment – but it was *situational priming* rather than the id.

Now the question was whether temporary accessibility from verbal priming could really do all of this. In 1973 I began to develop a way to study this issue with an undergraduate at Princeton, Carl Jones, and then continued to work on it with Steve Rholes when he arrived as a graduate student in 1974. The psycholinguistic roots of the project were apparent in the title of Carl Jones' senior thesis, “An experiment on the effect of language on nonlinguistic behavior.” The combining of cognitive psychology *and* social psychology was more evident in the final title of the published version of the research, “Category accessibility and impression formation” (see Higgins et al., 1977).

The first question was how to accomplish the verbal priming in a way that would increase a category's accessibility for more than just a few seconds, *and* in a manner

that would psychologically separate the verbal priming event itself from the target person information used to form an impression. We decided to use an *unrelated studies* paradigm. The participants were told that there were two separate studies conducted by two different experimenters that took place in two separate rooms (separated by a long hallway). Given our procedure, the increased accessibility from priming in “the first study” had to last several minutes to have an effect on judgment in “the second study.” I should note that some of these details do *not* appear in the published paper, but it influenced our decisions when designing subsequent priming studies. The participants were told that [p. 79 ↓] the purpose of the first study was “to examine the effects of information processing on perception,” and the purpose of the second study was to study “verbal comprehension.” It was within the first study that the verbal priming took place, and it was within the second study that participants were exposed to ambiguous target person information that they could use to form an impression of the target person.

In the supposed first study, the participants were told to name, as quickly as possible, the background color of ten different slides. Ostensibly to make the task more difficult, after naming the color of a slide the participants also had to repeat a word they had received aurally before the slide appeared. The verbal priming was accomplished by embedding the critical words among the other “memory” words, which meant that these primed words had to be remembered for several seconds. In this way, the verbal priming was strengthened, but the words themselves were secondary to the focal task. Indeed, by the time of the second study they did not consciously remember these “memory” words very well as episodic events.

There were four priming conditions involving four priming words and six other words:

- 1. *applicable, positive*: “adventurous,” “self-confident,” “independent,” “persistent”;
- 2. *applicable, negative*: “reckless,” “conceited,” “aloof,” “stubborn”;
- 3. *nonapplicable, positive*: “obedient,” “neat,” “satirical,” “grateful”;
- 4. *nonapplicable, negative*: “disrespect,” “listless,” “clumsy,” “sly.”

Note that in half of the conditions the priming words were “applicable” and in the other half they were “nonapplicable.” This was done to control for the possibility that somehow the positive or negative priming words would induce a positive or negative mood that,

in turn, would influence impressions. The “applicable” and “nonapplicable” words were selected to be equally positive or equally negative to control for this possibility. The verbal priming effect should occur only for the “applicable” words whose categorical meaning could be used to characterize the input information. They should *not* occur for the “nonapplicable” words. Only later did *applicability* become an important principle of knowledge activation theory in its own right. But this was the beginning.

Following the “perception” study, the participants went down the hallway and did the “reading comprehension” study. They were given a paragraph about a target person called “Donald” – who subsequently became infamous for overappearance in experimental studies – and they were told simply to familiarize themselves with the paragraph because later they would have to answer questions about it. It took subjects about two minutes to read the paragraph, which was composed of evaluatively ambiguous descriptions of Donald’s behaviors, such as the following *persistent-stubborn* description:

Once Donald made up his mind to do something it was as good as done no matter how long it might take or how difficult the going might be.
Only rarely did he change his mind even when it might well have been better if he had.

After the participants had read the paragraph, they were asked to characterize Donald’s behaviors. As predicted, the study found that in the “applicable” conditions, but not in the “nonapplicable” conditions, participants generally characterized Donald’s behaviors in terms of the positive categories in the positive verbal priming condition and in terms of the negative categories in the negative verbal priming condition. In order to check for awareness of the priming words, one set of participants were told that we wished to know “whether *anything* about the ‘perception’ study interfered with or affected their behavior in the ‘reading comprehension’ study, as it was still easy for us to change our procedure to avoid such problems.” In other words, they would *help* us by telling us about the priming words. Only one participant even noticed a connection between the priming words and how he characterized Donald.

[p. 80 ↓] The participants returned 10 to 14 days later for “another comprehension study.” (They were only debriefed after this second session.) Although the positive and

negative “applicable” participants during the first session had not differed in their ratings of Donald's desirability as a person, they did differ *now*, rating him as more desirable if they had been positively primed than if they had been negatively primed. Importantly, this effect on attitudes toward Donald was found both for those participants who had been asked explicitly in the first session to characterize Donald (overt characterization) *and* for those who had not received such explicit instructions (no overt characterization). For the latter participants, their attitudes toward Donald were spontaneous and evident two weeks after priming. Thus, for them, the priming effect on categorization was *covert* in the first session and had a lasting, if not increasing, effect on attitudes.

To my knowledge, this was the first study to demonstrate that verbal priming in one situation could influence how target information was categorized several minutes later in a totally separate situation, without any awareness of this happening. Moreover, this impact on categorization influenced attitudes toward the target two weeks later. Although I had predicted that this would happen, I was still surprised that it did. In fact, it was surprising to many other people as well and was considered a fluke by some (i.e., a type II error). It was not fully accepted until it was replicated several times by us and by others (e.g., Srull and Wyer, 1979, 1980). Of course, by now it has been replicated hundreds of times in many different forms in many different labs. It is no longer considered surprising. Thanks to Thomas Mussweiler, it was even given a thirtieth birthday party, the “Donald Symposium,” at the Society of Experimental Social Psychology in 2007.

I should say, however, that I was not just surprised that the priming really worked. I was disturbed by it. Indeed, I am still disturbed by it. Bruner's accessibility is not disturbing. It is reasonable and even beneficial for current input information to be captured by an accessible category because it is expected or because it satisfies some need. But it is neither reasonable nor beneficial for a category to be accessible simply because it happened to be, incidentally, verbally primed in a prior situation; and then determine, several minutes later, how a person's behavior is evaluatively categorized in a separate situation; and then, two weeks later, make that person seem more desirable or undesirable as a person. I will return to this problem at the end of this chapter.

Generalizing Priming and Accessibility Effects: The “Duncker Candle” Study

My first question after the “Donald” study was about the potential breadth of priming and accessibility effects. If priming could do this to people’s impressions of a target person, what else could it do? Working with Huttenlocher on “three-term series” problems had sparked my interest in problem-solving more generally. I became especially fascinated with “creativity” problems (e.g., Duncker, 1945). I wondered whether priming and accessibility could be used to get people to see something differently and thereby facilitate creative insight.

In 1976, I began working on this problem with William Chaires, another undergraduate at Princeton. To study this issue, we chose Duncker’s (1945) famous candle problem:

Subjects are seated at a table on which there is a cardboard wall, a candle, a full book of matches, and a box filled with thumbtacks.

Subjects are told that their task is to affix the candle to the cardboard wall so that the candle burns properly but does not drip wax on the table. The difficult part of the problem is to think of using the box as a platform for the candle, rather than just as a container for the tacks. The critical factor in solving the problem appears to be how subjects encode the box filled with thumbtacks.

[p. 81 ↓] What is difficult about this problem is that the most natural way for people to categorize the stimulus is *box of tacks*, which makes people think of the box as just a *container* for the tacks (see Glucksberg and Weisberg, 1966). The “of” categorization associates the box with the tacks in an *undifferentiated* manner, which highlights its function as a container. To solve the problem requires changing set and thinking of the box instead as a potential platform – a platform for the candle! If the stimulus were categorized as a “*box and tacks*,” this could *differentiate* the box from the tacks, establishing the box as more of an independent object.

The trick, then, would be to make the “and” construction temporarily more accessible than the “of” construction by priming the “and” construction. To accomplish this, we again used the “unrelated studies” paradigm. The participants were told that we wished to study the effects of interference on long-term memory; they would be shown a series of objects to remember and, in the “interference” condition, they would have to work on a problem before recalling the objects. The “interference” was working on the Duncker candle problem.

The “objects to be remembered” were various different objects, such as a banana, a football, a pair of scissors, plus some objects that were a container with content, such as a bowl containing cereal and a carton containing eggs. The experimenter labeled the objects when they appeared on a slide. In the “Of” priming condition, the slides with container-content objects were labeled as “bowl of cereal,” “carton of eggs,” and so on. In the “And” priming condition, these same slides were labeled as “bowl *and* cereal,” “carton *and* eggs,” and so on. After priming, all participants were given the instructions for the “interference” candle problem. There was a time limit of ten minutes to solve the problem.

In the “Of” priming condition, the average time working on the problem was nine minutes and only 20 percent solved the problem, which is quite typical for this problem. In the “And” condition, on the other hand, the average time was 4.5 minutes and 80 percent solved the problem – a dramatic improvement! These studies (Higgins and Chaires, 1980) demonstrated that priming and accessibility had a potential power over thought and behavior that was far greater than had been appreciated before. They were the first studies to demonstrate that priming could influence someone's behavior on a task, including creative behavior on a difficult problem-solving task. They showed that priming and accessibility effects are not just restricted to influencing people's recognition of *what things are* (e.g., someone's behavior is “persistent” versus “stubborn”); they can also influence people's recognition of *how things work*.

Beyond Accessibility from Priming

In 1978 I left Princeton for the University of Western Ontario in London, Ontario. For the first time I was in a department with colleagues who identified themselves as personality

psychologists, such as Dick Sorrentino. Up to this point I had thought of accessibility only in terms of priming and temporary accessibility. Now I began to consider the possibility of their being individual differences in the *chronic accessibility* of stored constructs that was separate from the temporary accessibility produced by priming. But there was no available measure of individual differences in chronic accessibility. How might such differences be measured?

Individual Differences in Chronic Accessibility

I wanted to distinguish accessibility from availability. Previous theories had emphasized individual differences in the kinds of constructs that are actually present in memory to be used to process the world, which involves a difference in construct *availability* [p. 82 ↓] (see Kelly, 1955; Markus, 1977; Mischel, 1973). In contrast, I wanted a measure of individual differences in construct *accessibility*. At Western, Gillian King and I ultimately developed a measure where we asked participants to list the traits (maximum of ten) for: (1) a type of person whom you like; (2) a type of person whom you dislike; (3) a type of person whom you seek out; (4) a type of person whom you avoid; and (5) a type of person whom you frequently encounter. There was a four- to five-minute delay between completing one question and receiving the next question – time that was filled with a nonverbal task to wipe out working memory. To capture the *accessibility* of a construct, the measure selected traits that were listed *first* in response to a question – *output primacy*.

Now that we had a measure of chronic accessibility, the question was whether individual differences in chronic accessibility influence how target person information is processed. To examine this question, we first gave participants our chronic accessibility measure. Then, in an “unrelated study” that took place about a week later, they read information about a target person. The essay with the target person information was individually tailored to each participant. Half of the trait-related behavioral descriptions exemplified one of that participant’s chronically accessible constructs, and the other half of the descriptions exemplified a trait that was not accessible to that participant but was an accessible construct for another participant in the study – a *quasi-yoking* design.

For example, the trait construct “friendly” might be a chronically accessible construct for one participant but not for another participant – although it was clearly available for the latter participant – and the essay description of the target person which *both* participants received would be as follows: “Person A is the kind of person who spontaneously strikes up conversations with others and who goes out of their way to say hello to someone.”

The first thing that we discovered was that individuals' chronically accessible constructs are surprisingly idiosyncratic (see Higgins et al., 1982). Looking at all possible pairs of participants, we found that the average percentage of overlap of accessible constructs was *less than 10 percent*, with more than half of the pairs having *no overlap at all* in their chronically accessible constructs! After reading the essay, there was a ten-minute delay filled with a nonverbal interference task, and then the participants were asked to reproduce the essay exactly, word for word (recall). Next they were asked to give their impression of the target person. For both participants' recall and impression of the target person, a trait description of the target person from the original essay was *less likely* to be deleted if it exemplified a participant's chronically accessible constructs than if it exemplified an inaccessible construct. There was also a high positive correlation between the evaluative tone of participants' impressions of the target and their *liking* for the target ($r = 0.70$), which means that individual differences in chronic accessibility can influence attitudes toward others as well.

Around the time that the new research on chronic accessibility was being conducted (1979), I went to the University of Michigan as a visiting professor. It was during my stay there when I began to develop a more general model of knowledge activation (see Higgins and King, 1981). In addition to distinguishing between available versus accessible constructs, chronic versus temporary accessibility, and applicable versus nonapplicable constructs, the model introduced two additional distinctions: active versus passive processing (i.e., controlled versus automatic); accessibility as “top of the storage bin” (what became Wyer and Srull’s “storage bin” model) versus accessibility as “energy potential” (what became my “synapse” model).

Active (Controlled) versus Passive (Automatic) Processing

In the early 1970s when I was a graduate student at Columbia, one of Michael Posner's [p. 83 ↓] collaborators, Robert Warren, came to Columbia as a new assistant professor. Posner and Warren (1972) had distinguished between *active*, conscious processes that are deliberate and controlled, versus *passive*, unconscious processes that occur automatically and are uncontrolled – a distinction that became more popularly known as *controlled* versus *automatic* processing (Shiffrin and Schneider, 1977). Whereas *set* is an active process in which conscious attention is deliberately directed toward the expected event, priming effects, such as those found in the Higgins et al. (1977) study, involve passive, automatic activation of constructs: they occur without intention and without conscious awareness (see Bargh and Pietromonaco, 1982; Higgins and King, 1981; Smith et al., 1992). This distinction was highlighted in Higgins and King (1981) for good reason – it is the fact that accessibility effects can occur without intention or awareness which makes them so intriguing and disturbing.

Accessibility as Energy Potential (The Synapse Model) versus Top of the Storage Bin

In Higgins and King (1981) I also addressed for the first time what it meant for something to be accessible. I probably could have continued doing research on accessibility without explicitly addressing this question. After all, I could do priming studies and measure chronic accessibility without having to make an explicit claim about what it meant for something to be accessible. But two events happened that motivated me to think about this basic issue. The first event was a visit to the University of Alberta in the spring of 1978 where I gave a talk about the "Donald" study. An animal learning psychologist listened carefully to the talk and then, during the discussion period, asked me – in a thoughtful and polite manner – what I meant by something

being “accessible.” It was an innocent, “pardon my ignorance” question that completely floored me. I had no idea how to answer his question, and simply said that I needed time to answer it and could we talk about it later – which we did. And it was true – I did need time to answer his question.

The second event was in the summer of 1978 at the Ontario Symposium on Social Cognition that took place in London, Ontario when I had a conversation with Bob Wyer. What we talked about – among many other things late into the night – was his new “storage bin” model of accessibility that he presented at the conference (Wyer and Srull, 1981). According to Wyer and Srull's storage bin model, priming a construct made it more accessible because when a construct is recently activated it is replaced on top of a layered storage bin, and constructs are subsequently used as a function of their position in the bin, beginning from the top of the bin. With time, other constructs are likely to be used and then deposited on top of the bin, thereby accounting for the decline of priming effects over time (i.e., accessibility decay). A construct that is frequently activated, that is, frequent priming, has a greater likelihood of having been activated recently and then deposited on top of the bin, thereby accounting for why decay effects over time are reduced when there is frequent priming (see Srull and Wyer, 1979, 1980).

Thanks to the Alberta question from months before, I had been thinking about the concept of “accessibility” for a while and I had a different metaphor in mind than a storage bin. My preferred metaphor was an energy cell whose energy or action potential is increased whenever the cell is activated, and this energy slowly dissipates with time. This was a more dynamic and less structural metaphor. Later on when I was at New York University (NYU), I realized that there were different possible “energy transmission” models, and my preference was a more biological system model than a battery-like model. I went to speak to a colleague of mine who was a neuroscientist, Tony Movshon. After explaining what I wanted my model to [p. 84 ↓] accomplish, he suggested that I read the work of Eric Kandel at Columbia (e.g., Kandel, 1976). I became convinced that the functioning of the “synapse” was my best metaphor, and the model became the *synapse model* (see Higgins et al., 1985). But, to return to my story, when I was talking to Bob in 1978, I had only a vague energy cell model. However, even then, it was different from Wyer and Srull's storage bin model.

An unusual thing happened in my conversation with Bob that night. We actually discussed what the essential differences between the two models were, and whether an experimental study could be designed to provide a critical test whose results would support one model but disconfirm the other. This was the first time when I had ever had such a conversation. It was the kind of conversation that philosophers of science claim that scientists should do – find ways to disconfirm a theory. Instead, most scientists look for ways to support their theory. But Bob and I tried to find an experiment that would disconfirm one of the two models. One critical difference between the models was that, according to the storage bin model, recent activation could maintain high accessibility, even more than an alternative frequently activated construct, as long as the recently activated construct (rather than the frequently activated construct) remained on top of the bin. The key was to make sure that no competing construct was activated during the delay period.

The full design and execution of such a critical test would take a few years – until after I left Western and went to NYU. At NYU I was blessed with a new colleague, John Bargh, who had very similar “burning issues” to me. Luckily for me, he was willing to collaborate, together with a graduate student Wendy Lombardi, to fully flesh out the design and deal with the issue of ensuring that no competing construct was activated following recent priming. The solution was to have participants perform a counting backward task immediately after the last prime, that is, the most recently activated construct. The positive and negative primes were the same trait terms used in Higgins et al. (1977), plus synonyms of those trait terms. For each positive-negative set of trait-related terms (e.g., persistent, determined, steadfast versus stubborn, obstinate, headstrong), either the positive or negative trait was primed most frequently and its opposite was primed only once but most recently (i.e., the last prime), *followed immediately by the counting backward task*. Thus, frequency of activation was pitted against recent activation and no competing construct was activated after recent priming.

Because there was no competing construct after the last prime, the storage bin model predicts that the construct primed most recently will dominate categorization of Donald's behaviors after both a short and a long delay filled with counting backward. But the synapse model makes a different prediction. After a very brief delay, the excitation level of the recently primed construct should be higher than that of opposite frequently primed construct because recent activation brings excitation to its maximal level and

the frequently primed construct has already decayed to some extent. This would result in Donald's behaviors being categorized in terms of the recently primed construct. But after a long two-minute delay, the frequently primed construct will have decayed less than the recently primed construct because, consistent with the synapse metaphor, the decay function for a construct that is multiply primed should be slower than the decay function for a construct primed only once. Given this, Donald's behaviors should now be categorized in terms of the frequently primed construct.

The synapse model, then, predicts a reversal over time: recent priming dominating categorization after a brief delay, and frequent priming dominating after a long delay. Again, unusual in science, I telephoned Bob Wyer, described the final design, and asked him whether he agreed with the competing predictions. He did, which made us happy and grateful, and we could then proceed with the study. The results of the study [p. 85 ↓] supported the predictions of the synapse model rather than the storage bin model (see Higgins et al., 1985), and this recency-frequency reversal effect over time was replicated by Lombardi et al. (1987). The same team later conducted a conceptual replication substituting chronic accessibility for frequent priming, still pitted against recent priming, and we replicated the reversal effect as a function of delay time (Bargh et al., 1988). This research program provided strong support for the synapse model, and it led to a revision of the storage bin model to account for this reversal effect (Wyer and Srull, 1989). It is also a fond memory that I will always cherish of collaborating with another scientist, Bob Wyer, on testing the competing predictions of two different theories.

Knowledge Activation Theory: The Next Generation

Explicitly distinguishing between accessibility from recent priming versus from frequent priming, and between accessibility from recent priming versus from chronic individual differences, turned out to be important not only for testing the unique predictions of the synapse model and the storage bin model, but also for appreciating a significant characteristic of accessibility – *its additive nature*. What the Higgins et al. (1985) and Bargh et al. (1988) studies demonstrated was that the accessibility of a construct lasted

longer when its source was frequent priming or chronic accessibility than when its source was just recent priming. Moreover, the comparable findings in these two sets of studies demonstrated that accessibility from chronic accessibility functioned like accessibility from frequent priming. Together, this suggests that combining different sources of accessibility would heighten accessibility and make it last longer. It was as if you could combine Bruner's expectancy source and his need source and make accessibility even stronger – something which had never been suggested. Was accessibility additive in this way? If it were, then people could not possibly know where a construct's current level of accessibility comes from – how much from chronic accessibility and how much from priming (an “aboutness” problem; see Higgins, 1998b).

Additivity in Accessibility

The storage bin model does not predict additivity of accessible sources. Instead, the accessibility of a construct from frequent priming or from chronic accessibility simply increases the likelihood that the construct will have been activated recently and then placed on top of the bin. That is, multiple sources simply increase the likelihood of obtaining an effect of recent priming. But there is evidence that accessibility *is* additive (Bargh et al., 1986; Higgins and Brendl, 1995). And this evidence includes a demonstration by Higgins and Brendl (1995) of an additivity effect on intensity of judgment from different levels of chronic accessibility combined with a short delay after recent priming plus prior frequent priming (two additional times) of the same recently primed construct. The condition of frequent priming plus recent priming with short delay should itself have placed the construct on top of the storage bin. From the perspective of the storage bin model, an additional factor of varying levels of chronic accessibility should not matter much, but in fact it mattered a lot.

Higgins and Brendl (1995) used a revised version of the standard measure of chronic accessibility described earlier which now took into account both primacy of output and frequency of output in order to distinguish different *levels of chronic accessibility* rather than just chronic versus nonchronic. Participants with varying levels of chronic accessibility for the construct “conceited” were (or were not) both frequently and recently primed with “conceited” and then immediately read about a target person “Sue” [p. 86 ↓] who displayed behaviors which pilot testing had shown were extremely

vague (i.e., they did *not* elicit spontaneous impressions that Sue was conceited). The participants were simply asked to give their impression of the target person. Like the pilot participants, even those experimental participants who were both recently and frequently primed and had only a short delay before judging Sue did *not* characterize her as being conceited. Nothing happened with this extremely vague input, with one exception. For those participants who had relatively high levels of chronic accessibility for the construct “conceited,” suddenly Sue appeared to be conceited. And the higher the participant's level of chronic accessibility for “conceited,” the *stronger* their impressions that Sue was conceited were. Compared to the other experimental participants and the pilot participants, it was as if these participants were *hallucinating* – they were seeing something that wasn't there. The strength of this additivity effect was, once again, disturbing in its implications, which I discuss later in the section on “Applicability to Social Issues.”

Assimilation versus Contrast Effects and the Principle of Judged Usability

The additivity effect means that it is not possible for anyone to know where the current level of a construct's accessibility is coming from. It could be from recent priming, frequent priming, chronic accessibility, or any combination of these. This makes it difficult – indeed, impossible – to know how much accessibility we should correct for if we believe that recent priming might bias our current judgments. What happens when people do try to correct for potential bias from priming? This became another major issue in the next generation of knowledge activation theory.

The phenomenon of correcting for bias from priming was first introduced by Martin (1986). What was highlighted early on was that the priming event itself was an episodic event that could be remembered later at the time that the judgment of the target information was to be made (see Lombardi et al., 1987). Because the priming event itself was irrelevant to what the target was like, it would be inappropriate for it to influence the judgment of the target. As I noted earlier, it is neither reasonable nor beneficial for incidental priming to affect judgment. Naturally, then, people would be motivated to correct for this possible source of bias. Unfortunately, it is impossible to

calibrate exactly how much accessibility to correct for and people often over correct. This produces a *contrast* effect where the target is not only not judged in terms of the primed construct but it is judged in terms of the *opposite*, competing construct, as when someone primed with “persistent” then judges Donald to be “stubborn.” Notice the correction is an overcorrection because the original ambiguous information is neither just persistent nor just stubborn.

The conditions that produce the standard assimilation effect of priming versus a contrast effect of priming have been studied extensively (for reviews, see Higgins, 1996; Mussweiler, 2003). Early on, knowledge activation theory emphasized the principle of *judged usability* to account for these effects (Higgins, 1996). Just because a construct is activated does not mean that it will necessarily be used subsequently to process input information. The construct could be judged to be irrelevant to use or judged to be inappropriate to use, as in the example above. Thus, the factors that determine whether a construct is activated must be distinguished from the factors that determine whether an activated construct is used or not.

Judged usability – which need not be a conscious process – can be a rather subtle process. The Higgins and Brendl (1995) study, for example, also had another condition with the usual ambiguous stimulus, in addition to the condition with the vague stimulus. For the ambiguous stimulus, the usual assimilation effect of priming was [p. 87 ↓] found when participants were not aware of the priming event. The question was, what would happen when participants were aware of the priming event? What the study found was that among these participants there was still an assimilation effect as a function of chronic accessibility – the stronger the chronic accessibility, the greater the assimilation effect. What this suggests is that when overall accessibility is higher than what would be expected from the priming event alone (i.e., higher because the level of accessibility derives from chronic accessibility and not just priming), then people are more likely to judge the accessible knowledge as being usable and less in need of correction even when they are aware of the priming event.

Knowledge Activation: Accessibility and Applicability Working Together

What determines whether stored knowledge is activated? According to knowledge activation theory, both the accessibility of stored knowledge and its applicability to an input determine whether stored knowledge is activated, where *applicability* refers to the amount of overlap between the features of stored knowledge and the attended features of the input. Accessibility and applicability work together to activate knowledge according to a *compensation rule*: the greater the accessibility of stored knowledge, the less applicability is needed for knowledge activation to occur, and the greater the applicability, the less accessibility is needed for knowledge activation to occur. The first half of this compensation rule is demonstrated in the Higgins and Brendl (1995) study described above in which strong chronic accessibility of the construct “conceited” increased the likelihood that the vague input, which was input for which the construct “conceited” had very low applicability, would be categorized in terms of the primed construct “conceited.” To appreciate the second half of the compensation rule, the difference between vague, ambiguous, and unambiguous input information needs to be appreciated (Higgins, 1996).

Vague versus ambiguous versus unambiguous input information concerns applicability. Input, such as a behavioral description, is vague when no stored construct has more than weak applicability to it. Input is ambiguous when there are least two alternative constructs with moderate to strong applicability to it – usually two alternatives with strong applicability in most past studies. Input is unambiguous when there is only one construct that is moderately or strongly applicable to it. To return to the second half of the compensation rule, with unambiguous input little accessibility is needed for knowledge activation because applicability is high and there are no competing alternatives. With ambiguous input, more accessibility of one of the alternatives is needed in order for that alternative to win over the competition. With vague input, much more accessibility is needed because the applicability is so low. Indeed, in the Higgins and Brendl (1995) study, even recent and frequent priming together were not enough to produce a categorization effect without the addition of chronic accessibility. For an

ambiguous input, in contrast, recent priming is sufficient to produce a categorization effect.

Knowledge Activation Theory: The Role of Motivation

When I began doing research on priming and accessibility, I thought of it as part of cognitive psychology. As I discussed earlier, it began as part of work that I was doing in psycholinguistics, and then later I extended it to impression formation. To me, this became a cognitive psychology/social psychology interface, the “social cognition” that I had been advised to pursue several years before. But this social cognition was grounded in [p. 88 ↓] cognitive psychology. It was part of the *information processlytizing* that I have described in an earlier paper with John Bargh (Higgins and Bargh, 1987). Perhaps the biggest change that has occurred for me over the last ten years – and, I believe, for others as well – is to rethink the *role of motivation* in knowledge activation principles.

Importance = Accessibility

Some relation between accessibility and importance has been recognized for a long time. For Bruner (1957), current needs such as hunger or poverty were postulated as increasing the accessibility of need-related objects such as food or coins (see also Higgins, 1981, for a discussion of motivation as a source of accessibility). In my self-discrepancy theory of how socialization creates strong ideal and ought self-guides in children (Higgins, 1989, 1991; see also my Regulatory Focus chapter in this volume), the motivational importance of a caretaker's response to a child's behavior was emphasized as a factor contributing to a child acquiring highly accessible (i.e., strong) ideal and ought self-guides. This developmental theory expanded the “accessibility = importance” proposal beyond current need satisfaction to trait-related constructs made important from past interactions with significant others (for later work building on this proposal, see Shah, 2003).

Later, this general “accessibility = importance” proposal provided the conceptual rationale for using individuals’ response times in reporting their personal ideal and ought goals as a measure of the relative importance of the promotion system (ideals) or the prevention system (oughts) in their self-regulation (Higgins et al., 1997). It was assumed, as before, that stored trait-related constructs which were important for self-regulation, such as being “friendly,” “athletic,” or “hardworking,” would have high chronic accessibility. Given this, it was now assumed that trait-related constructs that had high chronic accessibility (as measured by response times for reporting them) must be important for self-regulation. Subsequent studies testing the effects of promotion and prevention strength (i.e., the effects of highly accessible ideals and oughts) on emotions, decision making, and performance supported this assumption (see Higgins, 1998a; Higgins et al., 1997; see also my Regulatory Focus chapter in this volume).

Construct Activation from Priming = Action

During the same period that this “accessibility = importance” proposal was being developed and tested, another perspective on the role of motivation in knowledge activation was also being developed and tested. John Bargh, inspired by James’ (2007/1890) “ideo-motor action” notion, proposed that priming a stored construct would activate that construct and its associated behaviors, such as priming “elderly” would activate “walks slowly,” which would result in direct expression of the activated behavior. Indeed, in a now classic study, Bargh et al. (1996) did find that people walked slowly down a hallway after being primed with “elderly” – *in the absence of any elderly person*. (Priming effects on other kinds of behavior had been reported earlier [Carver et al., 1983; Higgins and Chaires, 1980], but under conditions where the current situation also contributed to what happened behaviorally.)

The “ideo-motor action” notion as proposed by James (2007/1890) is that activation of a bare idea can be sufficient to prompt action: “We think the act, and it is done” (see James, 2007/1890: 522). For James, the idea prompted action relatively immediately. For Bargh, there could be a delay between construct activation and carrying out a construct-related action. This greatly increases the significance of the ideo-motor action proposal. It is as if the construct activated by [p. 89 ↓] priming becomes an *intent* to directly express the construct in action. By becoming an intent, it is like Lewin’s “quasi-

need” where there is a tension created that cannot be reduced until an action fulfills the quasi-need (see Lewin, 1951). Thus, the tension can remain for a considerable period until the construct-related action is expressed.

This James-Bargh “ideo-motor action” notion of direct expression in action of an activated construct, where activation can occur from subliminal priming, inspired countless studies after the seminal Bargh et al. (1997) research. Construct activation from priming was shown to affect all kinds of behaviors, from aggressive behavior to achievement behavior to cooperative behavior (for reviews, see Bargh, 2005; Dijksterhuis and Bargh, 2001). What fascinated researchers was the implication that behavior could be determined by situational priming occurring outside of a person's awareness that activated a construct which then “needed” to be behaviorally expressed – “bypassing the will” (Bargh, 2005). The behavior was not directed by a person's preferences or desires or chronic goals but by construct activation itself, *by the activated idea itself*.

Construct Activation from Priming + Valence = Goal-Directed Action

Over the last five years or so, it has become increasingly clear that, generally speaking, construct activation from priming is *not* sufficient, by itself, to determine action. It is not the case that we think the act and it is done. A construct-related behavior activated by priming need not result in the direct expression of that behavior. Indeed, a behavior *opposite* to the primed behavior can be expressed depending on the valence of the activated construct.

To illustrate, consider a study by Cesario et al. (2006) which replicated the classic “elderly” study of Bargh et al. (1997). Cesario et al. (2006) argued that walking slowly would only make strategic sense for individuals who *like* the elderly because they would walk slowly in order to interact with them better. Some of the participants in the Cesario et al. (2006) study liked the elderly, and when the concept “elderly” was subliminally primed with pictures of elderly men, these participants afterward walked slowly down the hallway (in the absence of any elderly person) – just as Bargh et al. (1997) had

found. But for those participants who *disliked* the elderly, they walked *quickly* down the hallway after “elderly” was subliminally primed – as if they were trying to avoid interaction with an elderly person (even though there was, again, no elderly person present).

In earlier studies, Plaks and Higgins (2000) activated stereotypes for a teamwork partner that were associated with task performance, such as females being good on verbal tasks, and again found pragmatic behavioral choices rather than mimicry or direct expression of the stereotyped behavior; for example, *loafing* on a verbal task when preparing to work on that task with a female team partner. What these and the Cesario et al. (2006) studies show is that activation of categorical behavior, even from *subliminal* priming, does not necessarily result in direct expression of that behavior. Instead, people use the activated information to prepare for future interaction with a category member, and the behavior they prepare depends on a variety of factors, including their attitude toward category members (see Cesario et al., 2006 for a fuller discussion of this “motivated preparation to interact” proposal).

It seems, then, that activating a stored idea, even subliminally, does not necessarily produce direct expression of the activated behavior. It can produce the *opposite* behavior. Indeed, there is evidence that whether construct activation from priming even affects subsequent construct-related behavior can depend on whether value is attached to the activated construct. If we think of an activated construct as a state, such as the state of [p. 90 ↓] achievement or the state of cooperation, then attaching positive value to that state, either momentarily or chronically, transforms it into a *desired end-state*, that is, a goal to be fulfilled. If the current state of a person is discrepant from this desired end-state, then the person is motivated to take action to attain the desired end-state (see Custers and Aarts, 2005). Thus, the enacted behavior related to a primed construct is *behavior directed toward attaining a personal goal*. It becomes comparable to the behavior directed toward fulfilling personal ideals and oughts that have high chronic accessibility – chronically accessible goals that can unconsciously direct behavior (Higgins, 1998a).

There is now substantial evidence that behavioral effects of construct activation from priming vary as a function of the value (momentary or chronic) that is attached to the construct as an end-state (e.g., Ferguson and Bargh, 2004; for reviews, see Custers

and Aarts, 2005; Eitam and Higgins, 2010). There is evidence, for example, that participants who are deprived of water will drink more when the construct “drink” is primed than when it is not primed, but priming has no effect when participants are not water deprived (e.g., Strahan et al., 2002). There is also evidence that when subliminal priming increases the positivity of goal-supporting objects, it does so most strongly for individuals for whom the goal currently is more important (Ferguson, 2008).

Importance = Accessibility Revisited: Accessibility as Relevance

According to the traditional version of the “importance = accessibility” notion, when there is a goal (or need), goal-related constructs will be activated, which will make these constructs accessible. Construct accessibility depends on the frequency and recency of the activation and the time since the last activation. Motivation affects the frequency and recency of activation. It does not play a role, postactivation, in accessibility changes over time. Motivation, such as the need for accuracy or the need to avoid closure, can still affect judgment postactivation (e.g., Ford and Kruglanski, 1995), but this is through affecting the use of the accessible construct rather than the accessibility of the construct per se.

This traditional perspective on the relation between importance and accessibility did not change until very recently. Throughout the twentieth century, accessibility remained basically a cognitive variable that can be influenced by motivational factors. But more and more accessibility itself is becoming a motivational variable; or, more accurately, a *motivated cognition* variable. What is becoming increasingly clear is that accessibility depends on motivation. *Accessibility increases and decreases as a function of its relevance to current self-regulation.*

What determines the relevance of accessibility in current self-regulation? If accessibility levels serve current self-regulatory concerns, then we would expect the accessibility of goal-related constructs to be higher when a goal is higher in importance. Classic determinants of a goal's importance are the value of goal attainment and the likelihood of goal attainment. Does the accessibility of a goal-related construct increase as the

value and *likelihood* of goal attainment increase? Studies by Förster et al. (2005) indicate that indeed it does. Another classic determinant of a goal's importance is goal fulfillment or completion. The importance of a goal decreases after it is completed. Does the accessibility of a goal-related construct decrease after the goal is completed? There is evidence that this also happens (e.g., Förster et al., 2005). Notably, there is also evidence that goal completion *inhibits* the accessibility of goal-related constructs rather than just suppressing construct expression (see Liberman et al., 2007). The importance of a goal also decreases if it is fulfilled through a substitutable task (Lewin, 1951). Does the accessibility of a goal-related construct [p. 91 ↓] decrease after the goal is fulfilled through a substitutable task? There is evidence for this as well (Cesario et al., 2006).

The decrease in accessibility following goal completion which I described above can be understood in functional terms. Generally speaking, a completed goal that remains accessible could interfere with subsequent goal pursuit. The more important the goal, the more the risk of interference from goal-related constructs remaining accessible. This leads to the intriguing prediction that the more important a goal, the more goal-related constructs should decrease following goal completion. The Förster et al. (2005) studies supported this prediction as well.

But is it always the case that right after goal completion, the accessibility of goal-related constructs quickly decreases? Might such accessibility decay also depend on motivational relevance? Indeed, there is evidence for this from a study by Hedberg and Higgins (in press; see also Hedberg, 2007). Using the paradigm of Förster et al. (2005), the participants' task was to view a series of pictures and find when a picture of a pair of scissors was immediately followed by a picture of eyeglasses. Once this goal was completed, the accessibility of constructs related to "eyeglasses" was measured using a lexical decision task that appeared at different post-completion delay times. Hedberg and Higgins (in press) predicted that the accessibility decay function would be different for individuals with a strong promotion focus on accomplishment than for individuals with a strong prevention focus on security. In order to accomplish something new, promotion-focused individuals need to reduce the accessibility of "eyeglasses" in order that this old construct will not interfere with new accomplishments – precisely the kind of motivation discussed earlier. The more individuals are promotion-focused, the faster should be the rate of accessibility decay. Prevention-focused individuals, on the other hand, want to maintain the status quo until it is necessary to change. Thus, the

more individuals are prevention-focused, the *slower* should be the rate of accessibility decay – the opposite of promotion-focused individuals. These predicted *opposite* patterns of *accessibility decay* were found for promotion-focused and prevention-focused individuals.

The notion of accessibility as motivational relevance has been recently developed more fully and formally in a paper by Eitam and Higgins (2010). In a new framework called Relevance of a Representation (or ROAR for short), it is proposed that not all stimulated representations become active in the sense of being functionally available to mental processes such as categorization, planning, and effort allocation that impact judgment and behavior. ROAR proposes that a representation's impact on judgment and behavior over time does *not* derive from the maintenance of the representation's accessibility *per se*. Instead, the effects depend on the continuing relevance of the representation. Whereas the traditional notion is that priming produces accessibility that then decays at a certain rate as a function of factors like frequency of priming, ROAR suggests that what changes or is maintained is the motivational relevance of a representation. This means that the likelihood that stimulation of a representation would produce activation with judgmental and behavioral effects is independent of frequent or recent priming except for the effects that frequent or recent priming can have on motivational relevance.

Applicability to Social Issues

People's evaluations and decisions are influenced by the accessibility of their attitudes, beliefs, and past experiences (for reviews, see Förster and Liberman, 2007; Higgins, 1996). People's actions, from their voting decisions to their choice of activities, are influenced by the accessibility of action-related constructs (e.g., Fazio, 1989; Lau, 1989). This has important implications for social issues because the decisions and actions associated with accessible constructs [p. 92 ↓] include social discrimination and hostile actions toward others.

What is troubling about the relation between accessibility and action? If the current accessibility of a construct or goal for a person were related solely to its present utility for that person, or to its accuracy in representing the world that person is facing, then

there would not be a problem. What is troubling, however, is that this is not the case. There can be sources of accessibility that are independent of current utility or accuracy. This fact raises two separate troubling issues: the issue of *control* and the issue of *truth*. People want to be effective in making something happen – to be effective at control. People also want to be effective at establishing what's real, to be accurate, correct, and right – to be effective at truth (Higgins, 2011). The results of priming studies raise questions regarding people's effectiveness at control and at truth.

Recently, the control problem has received the most attention, inspired especially by the work of John Bargh and his collaborators. The issue is whether it is people themselves who make things happen, that is, who are in charge, or whether instead they are directed (without their intent or awareness) by situational priming of goals and constructs. This is clearly an important issue with significant implications for social issues, such as controlling behavioral discrimination and stereotyping (e.g., Devine and Monteith, 1999). It should be noted, however, that the above discussion of motivational factors underlying accessibility effects from priming offers some hope regarding the control problem. It suggests that individuals who are motivated, for example, to behave nonprejudicially could do so – even unconsciously. Indeed, there is research on individuals with chronically accessible egalitarian beliefs that supports this conclusion (Moskowitz et al., 1999).

If there is some hope for the control issue, what about the truth issue? The truth issue was always my “burning issue.” *Accessibility does not know where it comes from.* A current level of accessibility can be the result of any combination of recent situational priming (after a short delay or a long delay), frequent situational priming, and personal chronic accessibility. Knowledge activation itself is even more complex because all of these sources of accessibility combine with applicability to a target that can be ambiguous, vague, or relatively unambiguous. A person does not know where the activation comes from. It is impossible to calibrate accurately the contributions of these different sources. And the dominant tendency of people is to infer that the source of the activation is the contribution of the target, that is, that the activation is *about* the target's stimulus features (the “aboutness principle”; see Higgins, 1998b). The Higgins and Brendl (1995) study demonstrates this tendency to treat the target as the source of activation – to perceive Sue as “conceited” despite the vague input when there is strong accessibility. Poor Sue was hardly conceited, but she was judged as such under

conditions that maximized accessibility (i.e., frequent and recent priming with a short delay plus high chronic accessibility).

What does this say about the truth of our judgments and evaluations (and memories)? What does it mean for our ability to be accurate, correct, veridical (see also Kruglanski, 1989; Popper, 1959)? Does this false “aboutness” from uncalibrated sources of knowledge activation contribute to delusions and hallucinations, to false eyewitness testimony, to stereotypic and ethnocentric evaluations, to the conviction that others who disagree with our beliefs must be either lying or crazy? I believe that it does. It *is* a serious problem. And this problem cannot be solved by somehow ridding ourselves of the principles of knowledge activation. After all, the same principles underlie human education and learning. The knowledge activation principles are among the great trade-offs of being human. They necessarily produce errors of judgment, memory, decision making, and behavior, while at the same time they are a necessity if we are to learn from experience. [p. 93 ↓] They free us from the confines of the here-and-now, but it comes with a costly downside regarding truth (Higgins, 2000).

So what can we do about it? We can't surgically remove the functioning of knowledge activation principles without losing the benefits along with the costs. Perhaps this is where other people come in – the importance of not being an “island unto yourself.” We could, for example, compare our evaluation of a target to those of other people. This should improve validity, reliability, and objectivity because other people's sources of accessibility are likely to be different from ours while the target's actual properties remain constant (generally speaking). Indeed, all things being equal, this is probably a good way to reduce error. The fly in the ointment is that people tend to engage in such comparisons with in-group members – often a small number of significant others – with whom they want to create a *shared reality* (Echterhoff et al., 2009; Hardin and Higgins, 1996; Higgins, 1992; see also the Shared Reality chapter in this volume). By working for a shared reality, they bring their accessibility sources in line with one another, thereby undermining the benefits of having *independent* evaluations of a target. This is a strong argument in favor of having true *diversity* among individuals when a target is being evaluated or remembered. What's the bottom line? Introduce group diversity whenever possible to reduce the downside of knowledge activation principles on undermining the truth – increased benefits with reduced costs.

Diversity, then, is one solution to the truth problem created by accessibility and the other knowledge activation principles. But the truth problem has very broad significance that needs other kinds of solutions as well. The significance of the issue is evident in the important role that accessibility has been given in models of attitudes and behavior (e.g., Ajzen and Sexton, 1999; Fazio, 1990), interpersonal relations (e.g., Andersen and Chen, 2002), and motivated cognition more generally (e.g., Kruglanski et al., 2006). In the theory of planned behavior (see Ajzen and Sexton, 1999; see also the Planned Behavior chapter in this volume), for example, the functioning of each of the major sources of behavioral intentions (i.e., attitudes, norms, perceived control) depends on what is accessible. Thus, whether or not beneficial health behaviors, environmental behaviors, intergroup behaviors, and so on actually occur depends on what knowledge is accessible and activated when the behavioral choices are made. This has profound implications for interventions targeted at increasing beneficial behaviors. Methods need to be devised that increase the likelihood that stored knowledge which supports beneficial behavioral intentions is accessible when the choices are actually being made. Interventions targeted at accessibility could be as important for dealing with social issues as interventions targeted at any other contributing factor to behavioral intentions.

References

- Ajzen, I. and Sexton, J. (1999) Depth of processing, belief congruence, and attitude-behavior correspondence . In S. Chaiken, ed. and Y. Trope (eds), *Dual-Process Theories in Social Psychology* , pp. 117–138. New York: Guilford Press.
- Bargh, J.A. (2005) Bypassing the will: Toward demystifying the nonconscious control of social behavior . In R.R. Hassin, ed. , J.S. Uleman, ed. , and J.A. Bargh (eds), *The New Unconscious* , pp. 37–58.
- Bargh, J.A., Bond, R.N., Lombardi, W.J. and Tota, M.E. (1986) The additive nature of chronic and temporary sources of construct accessibility . *Journal of Personality and Social Psychology* , 50, 869–878. <http://dx.doi.org/10.1037/0022-3514.50.5.869>

Bargh, J.A., Chen, M. and Burrows, L. (1996) Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action . *Journal of Personality and Social Psychology* , 71, 230–244. <http://dx.doi.org/10.1037/0022-3514.71.2.230>

Bargh, J.A., Gollwitzer, P.M., Lee-Chai, A., Barndollar, K. and Trotschel, R. (2001) The automated will: Nonconscious activation and pursuit of behavioral goals . *Journal of Personality and Social Psychology* , 81, 1014–1027. <http://dx.doi.org/10.1037/0022-3514.81.6.1014>

Bargh, J.A., Lombardi, W.J. and Higgins, E.T. (1988) Automaticity of chronically accessible constructs in person × situation effects on person perception: It's just a matter of time . *Journal of Personality and Social Psychology* , 55, 599–605. <http://dx.doi.org/10.1037/0022-3514.55.4.599>

Bargh, J.A. and Pietromonaco, P. (1982) Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation . *Journal of Personality and Social Psychology* , 43, 437–449. <http://dx.doi.org/10.1037/0022-3514.43.3.437>

Bruner, J.S. (1957a) On perceptual readiness . *Psychological Review* , 64, 123–152. <http://dx.doi.org/10.1037/h0043805>

Bruner, J.S. (1957b) Going beyond the information given . In H. Gruber, ed. , G. Terrell, ed. , and M. Wertheimer (eds), *Contemporary Approaches to Cognition* . Cambridge, MA: Harvard University Press.

Carmichael, L., Hogan, H.P. and Walter, A.A. (1932) An experimental study of the effect of language on the reproduction of visually perceived form . *Journal of Experimental Psychology* , 15, 72–86. <http://dx.doi.org/10.1037/h0072671>

Carver, C.S., Gannellen, R.J., Froming, W.J. and Chambers, W. (1983) Modeling: An analysis in terms of category accessibility . *Journal of Experimental Social Psychology* , 19, 403–421. <http://dx.doi.org/10.1016/0022-1031%2883%2990019-7>

Cesario, J., Plaks, J.E. and Higgins, E.T. (2006) Automatic social behavior as motivated preparation to interact . *Journal of Personality and Social Psychology* , 90, 893–910.
<http://dx.doi.org/10.1037/0022-3514.90.6.893>

Chen, S. and Chaiken, S. (1999) The heuristic-systematic model in its broader context . In S. Chaiken, ed. and Y. Trope (eds), *Dual-process Theories in Social Psychology* pp. 73–96. New York: Guilford Press.

Custers, R. and Aarts, H. (2005) Beyond priming effects: The role of positive affect and discrepancies in implicit processes of motivation and goal pursuit . *European Review of Social Psychology* , 16, 257–300. <http://dx.doi.org/10.1080/10463280500435919>

Devine, P.G. and Monteith, M.J. (1999) Automaticity and control in stereotyping . In S. Chaiken, ed. and Y. Trope (eds), *Dual Processes* , pp. 339–360. New York: Guilford.

Dijksterhuis, A. and Bargh, J.A. (2001) The perception-behavior expressway: Automatic effects of social perception on social behavior . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 33, 1–40. San Diego: Academic Press.

Duncker, K. (1945) On problem solving . *Psychological Monographs* , 58, 5 (Whole No. 270).

Echterhoff, G., Higgins, E.T. and Levine, J.M. (2009) Shared reality: Experiencing commonality with others' inner states about the world . *Perspectives on Psychological Science* , 4, 496–521. <http://dx.doi.org/10.1111/j.1745-6924.2009.01161.x>

Eitam, B. and Higgins, E.T. (2010) Motivation in mental accessibility: Relevance Of A Representation (ROAR) as a new framework . *Social and Personality Psychology Compass* , 4, 951–967. <http://dx.doi.org/10.1111/j.1751-9004.2010.00309.x>

Fazio, R.H. (1989) On the power and functionality of attitudes: The role of attitude accessibility . In A.R. Pratkanis, ed. , S.J. Breckler, ed. , and A.G. Greenwald (eds), *Attitude Structure and Function* , pp. 153–179. Hillsdale, NJ: Erlbaum.

Fazio, R.H. (1990) Multiple processes by which attitudes guide behavior: The mode model as an integrative framework . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 23, 75–109. New York: Academic Press.

Ferguson, M.J. (2008) On becoming ready to pursue a goal you don't know you have: Effects of nonconscious goals on evaluative readiness . *Journal of Personality and Social Psychology* , 95, 1268–1294. <http://dx.doi.org/10.1037/a0013263>

Ferguson, M.J. and Bargh, J.A. (2004) Liking is for doing: Effects of goal pursuit on automatic evaluation . *Journal of Personality and Social Psychology* , 88, 557–572. <http://dx.doi.org/10.1037/0022-3514.87.5.557>

Ford, T.E. and Kruglanski, A.W. (1995) Effects of epistemic motivations on the use of accessible constructs in social judgment . *Personality and Social Psychology Bulletin* , 21, 950–962. <http://dx.doi.org/10.1177/0146167295219009>

Förster, J. and Liberman, N. (2007) Knowledge activation . In A.W. Kruglanski, ed. and E.T. Higgins (eds), *Social Psychology. Handbook of Basic Principles* (2nd edition), pp. 201–231. New York: Guilford.

Förster, J., Liberman, N. and Higgins, E.T. (2005) Accessibility from active and fulfilled goals . *Journal of Experimental Social Psychology* , 41, 220–239.

Glucksberg, S. and Weisberg, R.W. (1966) Verbal behavior and problem solving: Some effects of labeling in a functional fixedness problem . *Journal of Experimental Psychology* , 71, 659–664. <http://dx.doi.org/10.1037/h0023118>

Hedberg, P.H. (2007) What remains on your mind? Motivated up and down regulation of accessible knowledge . PhD dissertation, Columbia University.

Hedberg, P.H. and Higgins, E.T. (in press) What remains on your mind after you are done? Flexible regulation of knowledge accessibility . *Journal of Experimental Social Psychology* .

Higgins, E.T. (1989) Continuities and discontinuities in self-regulatory and self-evaluative processes: A developmental theory relating self and affect . *Journal of Personality* , 57, 407–444. <http://dx.doi.org/10.1111/j.1467-6494.1989.tb00488.x>

Higgins, E.T. (1991) Development of self-regulatory and self-evaluative processes: Costs, benefits, and trade offs . In M.R. Gunnar, ed. and A. Sroufe (eds), *Minnesota Symposium on Child Psychology* , pp. 125–165. Hillsdale, NJ: Erlbaum.

Higgins, E.T. (1996) Knowledge activation: Accessibility, applicability, and salience . In E.T. Higgins, ed. and A.W. Kruglanski (eds), *Social Psychology. Handbook of Basic Principles* , pp. 133–168. New York: Guilford.

Higgins, E.T. (1998a) Promotion and prevention: Regulatory focus as a motivational principle . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 30, 1–46. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960381-0>

Higgins, E.T. (1998b) The aboutness principle: A pervasive influence on human inference . *Social Cognition* , 16, 173–198. <http://dx.doi.org/10.1521/soco.1998.16.1.173>

Higgins, E.T. (2000) Social cognition: Learning about what matters in the social world . *European Journal of Social Psychology* , 30, 3–39. <http://dx.doi.org/10.1002/%28SICI%291099-0992%28200001/02%2930:1%3C3::AID-EJSP987%3E3.0.CO;2-I>

Higgins, E.T. (in press) Beyond Pleasure and Pain: How Motivation Works . New York: Oxford University Press.

Higgins, E.T. and Bargh, J.A. (1987) Social cognition and social perception . *Annual Review of Psychology* , 38, 369–425. <http://dx.doi.org/10.1146/annurev.ps.38.020187.002101>

Higgins, E.T., Bargh, J.A. and Lombardi, W. (1985) The nature of priming effects on categorization . *Journal of Experimental Psychology. Learning, Memory and Cognition* , 11, 59–69. <http://dx.doi.org/10.1037/0278-7393.11.1.59>

Higgins, E.T. and Brendl, M. (1995) Accessibility and applicability: Some 'activation rules' influencing judgment . *Journal of Experimental Social Psychology* , 31, 218–243. <http://dx.doi.org/10.1006/jesp.1995.1011>

Higgins, E.T. and Chaires, W.M. (1980) Accessibility of interrelational constructs: Implications for stimulus encoding and creativity . *Journal of Experimental Social Psychology* , 16, 348–361. <http://dx.doi.org/10.1016/0022-1031%2880%2990027-X>

Higgins, E.T. and King, G. (1981) Accessibility of social constructs: Information processing consequences of individual and contextual variability . In N. Cantor, ed. and J. Kihlstrom (eds), *Personality, Cognition, and Social Interaction* , pp. 69–121. Hillsdale, NJ: Lawrence Erlbaum Associates.

Higgins, E.T., King, G.A. and Mavin, G.H. (1982) Individual construct accessibility and subjective impressions and recall . *Journal of Personality and Social Psychology* , 43, 35–47. <http://dx.doi.org/10.1037/0022-3514.43.1.35>

Higgins, E.T., Rholes, W.S. and Jones, C.R. (1977) Category accessibility and impression formation . *Journal of Experimental Social Psychology* , 13, 141–154. <http://dx.doi.org/10.1016/S0022-1031%2877%2980007-3>

Higgins, E.T., Shah, J. and Friedman, R. (1997) Emotional responses to goal attainment: Strength of regulatory focus as moderator . *Journal of Personality and Social Psychology* , 72, 515–525. <http://dx.doi.org/10.1037/0022-3514.72.3.515>

Huttenlocher, J. and Higgins, E.T. (1971) Adjectives, comparatives, and syllogisms . *Psychological Review* , 78, 487–504. <http://dx.doi.org/10.1037/h0031858>

James, W. (2007) *The Principles of Psychology*, Vol. 2 . New York: Cosimo (Original work published 1890).

Kandel, E.R. (1976) *Cellular Basis of Behavior. An Introduction to Neurobiology* San Francisco: W.H. Freeman.

Kelley, H.H. (1950) The warm-cold variable in first impressions of persons . *Journal of Personality* , 18, 431–439. <http://dx.doi.org/10.1111/j.1467-6494.1950.tb01260.x>

Kelly, G.A. (1955) *The Psychology of Personal Constructs* . New York: W.W. Norton.

Kruglanski, A.W. (1989) Lay Epistemics and Human Knowledge: Cognitive and Motivational Bases . New York: Plenum.

Kruglanski, A.W., Erbs, H-P., Pierro, A., Mannetti, L. and Chun, W.Y. (2006) On parametric continuities in the world of binary either ors . *Psychological Inquiry* , 17, 153–165. http://dx.doi.org/10.1207/s15327965pli1703_1

Lau, R.R. (1989) Construct accessibility and electoral choice . *Political Behavior* , 11, 5–32. <http://dx.doi.org/10.1007/BF00993365>

Lewin, K. (1951) *Field Theory in Social Science* . New York: Harper.

Liberman, N., Förster, J. and Higgins, E.T. (2007) Completed versus interrupted priming: Reduced accessibility from post-fulfillment inhibition . *Journal of Experimental Social Psychology* , 43, 258–264. <http://dx.doi.org/10.1016/j.jesp.2006.01.006>

Lombardi, W.J., Higgins, E.T. and Bargh, J.A. (1987) The role of consciousness in priming effects on categorization . *Personality and Social Psychology Bulletin* , 13, 411–429. <http://dx.doi.org/10.1177/0146167287133009>

Markus, H. (1977) Self-schemata and processing information about the self . *Journal of Personality and Social Psychology* , 35, 63–78. <http://dx.doi.org/10.1037/0022-3514.35.2.63>

Martin, L.L. (1986) Set/reset: Use and disuse of concepts in impression formation . *Journal of Personality and Social Psychology* , 51, 493–504. <http://dx.doi.org/10.1037/0022-3514.51.3.493>

Mischel, W. (1973) Toward a cognitive social learning reconceptualization of personality . *Psychological Review* , 80, 252–283. <http://dx.doi.org/10.1037/h0035002>

Moskowitz, G.B. Gollwitzer, P.M., Wasel, W. and Schaal, B. (1999) Preconscious control of stereotype activation through chronic egalitarian goals . *Journal of Personality and Social Psychology* , 77, 167–184. <http://dx.doi.org/10.1037/0022-3514.77.1.167>

Mussweiler, T. (2003) Comparison processes in social judgment: Mechanisms and consequences . *Psychological Review* , 110, 472–489. <http://dx.doi.org/10.1037/0033-295X.110.3.472>

Plaks, J.E. and Higgins, E.T. (2000) The pragmatic use of stereotypes in teamwork: Motivational tuning to inferred partner/situation fit . *Journal of Personality and Social Psychology* , 79, 962–974. <http://dx.doi.org/10.1037/0022-3514.79.6.962>

Popper, K.R. (1959) *The Logic of Scientific Discovery* . New York: Basic Books. (Original work published 1934).

Posner, M.I. and Warren, R.E. (1972) Traces, concepts, and conscious constructions . In A.W. Melton, ed. and E. Martin (eds), *Coding Processes in Human Memory* . Washington, DC: Winston.

Sedikides, C. and Skowronski, J.J. (1991) The law of cognitive structure activation . *Psychological Inquiry* , 2, 169–184. http://dx.doi.org/10.1207/s15327965pli0202_18

Shah, J. (2003) The motivational looking glass: How significant others implicitly affect goal appraisals . *Journal of Personality and Social Psychology* , 85, 424–439. <http://dx.doi.org/10.1037/0022-3514.85.3.424>

Shiffrin, R.M. and Schneider, W. (1977) Controlled and automatic human information processing: II. Preceptual learning, automatic attending, and a general theory . *Psychological Review* , 84, 127–190. <http://dx.doi.org/10.1037/0033-295X.84.2.127>

Smith, E.R., Stewart, T.L. and Buttram, R.T. (1992) Inferring a trait from a behavior has long-term, highly specific effects . *Journal of Personality and Social Psychology* , 62, 753–759. <http://dx.doi.org/10.1037/0022-3514.62.5.753>

Srull, T.K. and Wyer, R.S. (1979) The role of category accessibility in the interpretation of information about persons: Some determinants and implications . *Journal of Personality and Social Psychology* , 37, 1660–1672. <http://dx.doi.org/10.1037/0022-3514.37.10.1660>

Srull, T.K. and Wyer, R.S. (1980) Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgments . *Journal of Personality and Social Psychology* , 38, 841–856. <http://dx.doi.org/10.1037/0022-3514.38.6.841>

Strahan, E.J., Spencer, S.J. and Zanna, M.P. (2002) Subliminal priming and persuasion: Striking when the iron is hot . *Journal of Experimental Social Psychology* , 38, 556–568. <http://dx.doi.org/10.1016/S0022-1031%2802%2900502-4>

Wyer, R.S. and Srull, T.K. (1981) Category accessibility: Some theoretical and empirical issues concerning the processing of social stimulus information . In E.T. Higgins, ed. , C.P. Herman, ed. , and M.P. Zanna (eds), *Social Cognition: The Ontario Symposium* , 1, 161–197. Hillsdale, NJ: Erlbaum.

Wyer, R.S. and Srull, T.K. (1989) *Memory and Cognition in Its Social Context* . Hillsdale, NJ: Erlbaum.

<http://dx.doi.org/10.4135/9781446249215.n5>

Handbook of Theories of Social Psychology: Volume 1

A Theory of Impulse and Reflection

Contributors: Fritz Strack & Roland Deutsch

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "A Theory of Impulse and Reflection"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n6>

Print pages: 97-118

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n6>

[p. 97 ↓]

Chapter 5: A Theory of Impulse and Reflection

Fritz Strack, ed. and Roland Deutsch, ed.

Abstract

The aim of the present chapter is to outline our theory of the dual determination of human behavior (Strack and Deutsch, 2004) in a larger context. Specifically, we attempt to demonstrate that impulse and reflection have long been identified as antagonistic forces. Although this distinction has emerged in various cultural and religious contexts as well as in different scientific disciplines, there exists no coherent account that describes the interaction of these influences along a sequence of information processing. Thus, it seemed justified to build a Reflective-Impulsive Model (RIM) that links thinking and feeling with behavior and allows connecting different research traditions and applied phenomena with the described duality. We begin by laying out the larger context in which we see the model to be embedded.

The Dual Determination of Behavior

Cultural Manifestations

In Greek philosophy, Plato created an allegory in his Phaedrus (Rowe, 1998) in which he used two horses and a chariot to illustrate the different underlying mechanisms. Specifically, the chariot is meant to symbolize the human soul (in love) that is jointly

pulled forward by both a bright and a dark horse. The two horses differ in their personality such that the bright horse is temperate, obedient and in no need of the whip while the dark horse acts tempestuously and does not yield to whip or spur. However, despite the dark horse's predominantly negative characteristics, Plato assigns it a crucial role in spontaneously approaching the other person and experiencing the joys of love. Thus, the lesson of this allegory seems to be that both modes of operation are necessary under different circumstances and that impulsive mechanisms are important for emotional action.

Quite a different lesson comes from the Christian religion, which has made the two modes of behavior a central theme. In particular, the Christian doctrine (e.g., Sorabji, 2002) holds that virtuous behavior may be in conflict with temptations. In this case, the "spirit" may be willing to guide one's action [p. 98 ↓] toward a positively evaluated end, the "flesh," however, may be weak and cause the person to engage in sinful behavior. Despite its negative evaluation, the Christian religion acknowledges that "sinning" (i.e., deviating from the path of virtue) will repeatedly manifest itself in human behavior. Therefore, it is suggested that people purge themselves from the sins by verbalizing and repenting their sinful behaviors. Thus, the Christian opposition of spirit and flesh explicitly acknowledges the duality of human behavior and recognizes the emergence of conflicts. Moreover, it capitalizes on their inevitability and suggests solutions that allow actors to remain good despite their occasionally bad behaviors. It is important to note that the Christian religion subscribes to a philosophically dualistic doctrine in which virtuous and more reflective processes are seen as mental or spiritual phenomena whereas sinful and more impulsive mechanisms are assumed to reside in humans' physical endowment.¹

Finally, most legal systems make a distinction between criminal actions that have been conducted with the intention to achieve a given outcome and those that are the consequence of mere negligence or of affective influences. Because punishment is determined by a person's guilt and guilt is related to the intentionality of sinning, premeditated crimes are believed to require a harsher punishment than the other two types of criminal behavior. This differentiation also seems to be based on assumptions about different determinations of behavior, such that more responsibility is assigned to one than to the other. Responsibility, in turn, is related to reflective processes where the

anticipated outcome determines the action. A lack of responsibility is more likely to be found for impulsive behaviors.

These examples demonstrate that the idea of a dual determination of human behavior is part of Western cultural knowledge that manifests itself in different ways. Of course, this is more a social acknowledgment of the existing duality than a precise analysis of its structure and dynamics.

Personality Differences

As for any human characteristic, the question arises if, how and why humans differ on a relevant dimension. Not surprisingly, the distinction between reflective and impulsive processes has long been the focus of psychological theorizing.

One of the earliest and best-known approaches that focus on the described duality is psychodynamic theory (Freud, 1927). Specifically, impulses were assumed to originate from the id and to seek immediate behavioral expression. The ego, however, has the function to restrain the impulses and to reconcile their behavioral consequences with the requirement of the situation. While the id is governed by the pleasure principle, the ego obeys the reality principle. These two instances of the “psychic apparatus” along with the superego that contains social norms are assumed to evolve as a function of early socialization. Obviously, the id is in charge of expressing behavioral impulses whereas the ego reflects about their situational and normative appropriateness.

The idea of a psychic instance that controls impulses is also part of more recent theories of personality (e.g., Carver, 2005). Under the label of “ego psychology,” Block and Block (1980) have emphasized ego control as an important variable. High ego control can be understood as the dominance of reflective tendencies, low ego control as that of impulsive reactions. Not surprisingly, individuals who are high on this dimension are more likely to delay gratification (Funder and Block, 1989).

Because a moderate level of ego control is often considered to be most desirable, it would imply an adaptive balance between constraint and hedonistic enjoyment. Empirically, this claim has been supported by the outcomes of research on ego

resilience, which describes the flexibility with which a person reacts to requirements of a situation in terms of more reflective or more impulsive responses. Interestingly, individuals found to be high on ego resilience had [p. 99 ↓] moderate levels of ego control (Asendorpf and van Aken, 1999).

Within the more conventional trait approaches, the five-factor model (Costa and McCrae, 2009) plays a central role. In this framework, the balance between reflection and impulse is best represented by the factor called “conscientiousness,” which juxtaposes careful and planned deliberation with rash and unorganized reactions. Also, the factor “agreeableness” may be somewhat related to the reflective-impulsive dimension (see Carver, 2005) in that it stands for overcoming egotistic tendencies.

Interestingly, the dimensions of “ego control” on the one hand and “conscientiousness” and “agreeableness” on the other seem to be empirically related. That is, people who were low in ego control also had low scores in both conscientiousness and agreeableness (Asendorpf and van Aken, 1999).

Another scale to be mentioned comes from Zuckerman's (1994a) work on sensation seeking. He identified a factor labeled “impulsive unsocialized sensation seeking” (IUSS) that plays a role in antisocial personality disorders (Zuckerman, 1994b) and is negatively correlated with conscientiousness and agreeableness (Zuckerman, 1996)

Within the five-factor model, Whiteside and Lynam (2001) have identified four distinct personality facets that are associated with impulsive behavior. They included the presence of both urgency and sensation seeking and the absence of both premeditation and perseverance.

Theoretical Approaches

Approaches from Economics

A central feature of impulsive responses is that they are not future-oriented. Specifically, impulsive characteristics or processes are understood to cause behaviors to be immediately expressed whereas their reflective counterparts allow for temporal postponement (McClure et al., 2004). This dimension of temporal delay is deeply rooted in economic thinking where the immediate consumption stands in conflict with saving, investment and consumption at a later time (Loewenstein and Thaler, 1997, 2005). In economics, this topic has come up under the name of “intertemporal choice” and has long been explained by the concept of “temporal discounting” (e.g., Ainslie, 2005). This idea implies that individuals differ in the degree to which a given temporal distance decreases the utility of a given outcome. To compensate for this devaluation, the delayed outcome must be increased by the same degree. This rule, however, was oversimplified and the discount rate was shown to vary within one person. For example, the same compensation that would be necessary to get people to accept a given delay that would start from the present would be lower if the same delay would begin ten years from now. This hyperbolic discounting was seen as a way of explaining impulsive behavior (Ainslie, 1975) and a means of describing it. More recently, however, behavioral economists have abandoned the hope of fully understanding the delay of reward by identifying the right discount rate or function. Instead, insights from psychology are believed to provide some important progress toward an understanding of the underlying processes (Frederick et al., 2003).

The Role of Impulse in Social Psychology

Early Approaches

It is interesting to note that in the first systematic approach to social psychology, social behavior was explained to be impulsively determined. In 1895, the French medical doctor Gustave Le Bon (1895) published a book titled *The Crowd* in which individuals were described as slaves of impulses if their actions occurred in a social context. When by themselves, their behavior was guided by [p. 100 ↓] reason and responsibility. However, while in groups, they were assumed to lose their “conscious personality” and be suggestively influenced by the behavior of others. Resisting these suggestive impulses requires a “strong personality”. Otherwise, social behavior was seen to be driven by impulsivity, irritability, a lack of critical judgment and logical thinking and an exaggeration of emotions. According to Le Bon, these characteristics were localized in a “group mind” that was seen to be in charge of emotional behaviors. Thus, social action was understood not to be a function of its anticipated consequences but the result of uncontrolled forces from outside the individual. Moreover, the internal mechanisms that are elicited in social settings are assumed to differ from those that are operating under individual conditions. Thus, Le Bon's analysis has invoked two systems with different operating principles that determine behavior under different circumstances. The two minds that he had been identified may be understood as two sets of processes of which either one or the other was operating. As a consequence, the author's theory does not imply the possibility of a conflict between the impulsive “group mind” and its reflective “individual” counterpart.

Such a conflict, however, was explicitly admitted by Sigmund Freud (1993/1921) who has criticized Le Bon's notion of a group mind. While denying the society as the source of impulsiveness, Freud acknowledged the existence of different operating principles which he identified *within* the individual. Specifically, the more reflective ego is often seen to be in conflict with the more impulsive id that is mainly located in the unconscious. Unlike in Le Bon's theory, human thinking, feeling and acting was believed to be the product of the joint operation of the two instances. However, both theories

agree that behavior is not only determined by rational considerations about its outcome but also by emotional forces. Also, both approaches agree that affective impulses may undermine a more reflective control of behavior.

The Re-Emergence of Impulse in Social Psychology

During the last 30 years, the notion of impulse in social psychology has re-emerged in various forms (Strack and Deutsch, 2007), perhaps the most important ones being the distinction between automatic and controlled processes (e.g., Bargh, 1994; Schneider and Shiffrin, 1977), the distinction between rules and associations (e.g., Sloman, 1996), and the distinction between heuristic and systematic processing (e.g., Chaiken, 1980; Tversky and Kahneman, 1974). The belief that there is more than one way in which people can think, feel and act has paved the way to a view of human behavior that is not exclusively guided by the anticipation of its outcomes.

From a social psychological perspective, various duality theories have been summarized in the contributions to a book by Chaiken and Trope (1999). Instead of reviewing each of them, we shall merely point at some landmark theories and findings that we think have had a particularly strong influence (see also Deutsch and Strack, 2006a). However, it is important to note that notions of duality have been prevalent not only in social psychology, but also in other subfields of psychology. In cognitive psychology, for example, there has been a discussion about the possibility that the world may not only be represented in linguistically defined concepts but also in an analogue fashion. In particular, it was argued (e.g., Kosslyn, 1980) and demonstrated (e.g., Farah, 1988) that perceptual representations may play an important role alongside their conceptual counterparts. This dual representation implies that people may simultaneously possess a perceptual experience on the one hand and knowledge or beliefs about what is the case on the other (Sloman, 1996; Strack, 1992). Typically, the two representations are in line with one another. At times, however, they may be inconsistent. Extending findings from judgmental heuristics (Greifeneder and Bless, 2007), one would suspect that motivation and cognitive capacity determine if the [p. 101]

↓] experiential or the noetic representation will dominate further processing. That is, the knowledge may override the experience.

Visual illusions may serve as an example (cf. Sloman, 1996). Being asked to judge the length of the two lines in the Müller-Lyer illusion, people may perceive the line between the open arrows to be shorter than that between the closed ones. At the same time, they may know that this impression is not caused by the actual length of the line but by contextual circumstances. As a consequence, they may not base their judgment solely on their perceptual experience but may use their knowledge to correct it (Strack, 1992; Koriat and Levy-Sadot, 1999; Sloman, 1996), which suggests different processes are connected with different contents.

The same is true for internally elicited experiences such as feelings. As Schwarz and Clore (e.g., Schwarz and Clore, 1983, 1996) have repeatedly demonstrated, feelings may enter judgments "as information." This means the effects that subjective experiences exert on judgments may be qualified by knowledge about their origin. Thus, positive affect that comes, for example, from the current weather conditions will not be used as a basis for judgments of happiness with one's life as soon as the person making the judgment becomes aware of the source. The same reasoning applies to "cognitive feelings" that are not accompanied by strong affect. Most prominently, the impact of feelings of fluency may be qualified if they are not representative for the judgments but originate from some irrelevant facet of the task or its context (e.g., Schwarz et al., 1991). Again, it should be noted that the described correction requires cognitive capacity and is less likely to occur under suboptimal circumstances (Strack et al., 1982).

Duality in Stereotypes and Attitudes

The impact of duality models on modern psychology was mediated by its influence in the domain of attitudes. At first, duality models had focused on the processes suggested as routes to persuasion. In the spirit of Tversky and Kahneman's (1974; Kahneman and Tversky, 1982) notion of heuristics as ways of simplifying judgments under conditions of uncertainty, the process of persuasion was increasingly understood to be based on two different types of psychological processes (e.g., Chaiken, 1980; Petty and

Cacioppo, 1986). On the one hand, the systematic route to persuasion will draw on the strength of arguments to come to a response to a persuasive message. Taking this more effortful route would require the appropriate motivation and cognitive capacity. On the other hand, the heuristic route to attitude change does not require scrutinizing the quality of the arguments but merely a less effortful check of peripheral cues such as the attractiveness of the communicator or the mere length of the persuasive message. Variants of these basic strategies and their consequences have been elaborated in different theories of attitude change. However, it has also been argued that the different routes to persuasion are not distinct in their psychological processes but in the type of information that enters into the judgments (Kruglanski and Thompson, 1999).

In a seminal paper, Devine (1989) offered a new perspective on prejudice by distinguishing between automatic and controlled processes in stereotyping. Specifically, she observed that although high and low prejudiced people may differ in their stereotyped beliefs about characteristics of group members, they did not differ in the accessibility of stereotypic contents in memory. Instead, low prejudice individuals seemed to be more motivated to prevent these activated thoughts from influencing judgment and action – at least if motivation and capacity to control prejudice are high (cf. Dunton and Fazio, 1997). These findings provide a strong hint that prejudice may be represented in two ways, as an evaluative belief and as a habitual tendency to activate certain thoughts.

The distinction between automatic and controlled components has also been applied to attitudes in general. Specifically, it has been argued that evaluations may exist not [p. 102 ↓] only as momentary beliefs about something being good or bad, which may be constructed on the spot by reflection, but also as over-learned object-evaluation associations in memory (e.g., Fazio et al., 1986; Greenwald and Banaji, 1995; Wilson et al., 2000). Such evaluative associations may be activated automatically when relevant objects are perceived and may then immediately trigger behavioral tendencies (cf. Chen and Bargh, 1999; Solarz, 1960). Among these behavioral manifestations are the facilitations and inhibitions of reactions that are either compatible or incompatible with a given valence.

This reasoning has led to a duality in the domain of attitudes that has exerted an enormous influence on thinking in social psychology. It has caused the distinction

between explicit and implicit attitudes (Greenwald and Banaji, 1995; Wilson et al., 2000; Wittenbrink and Schwarz, 2007), and has elicited a methodological discussion on the various procedures of measuring the implicit variant (De Houwer and Moors, 2007). Some procedures of “indirect” assessments of attitudes are based on pre-existing reactions that signal valence (like a smile or a frown), which are more easily executed in response to stimuli of compatible valence (Dimberg et al., 2002; Neumann et al., 2005). The response speed is thus an index of evaluative processes, which can be mapped even when participants are instructed to respond to a criterion other than valence such as the mere occurrence of the stimulus (Neumann et al., 2005). Other indirect attitude measures do not rely on pre-existing valence-behavior links, but create these links experimentally. For example, participants in Fazio and colleagues' evaluative priming task (Fazio et al., 1986) are trained to categorize clearly positive and negative words according to their valence by making previously neutral responses (i.e., key presses). In a later priming phase, these responses are facilitated when the target stimuli are preceded by evaluatively congruent prime stimuli. In a similar vein, the Implicit Association Test (IAT) (Greenwald et al., 1998) requires participants to categorize targets toward which the participants are suspected to have an implicit attitude with two keys. These keys, however, are also used to categorize clearly positive or negative stimuli, and after some training, the evaluation trials occur intermixed with the categorization trials. Depending on conditions, the category of interest shares the key with positive or negative stimuli. Depending on whether categorizations are quicker when the concept is mapped on the positive or negative key, the target concept is assumed to trigger a positive or negative automatic evaluation. For example, if participants are faster at pressing a particular key to categorize people as “old” if the same key is also used to indicate objects with a negative (compared to a positive) valence then this person is attributed a negative implicit attitude toward the elderly. Despite some controversies about the concept of implicit attitudes and the techniques for their assessment (e.g., Blanton and Jaccard, 2006; De Houwer et al., 2009; Fiedler et al., 2006; Rothermund and Wentura, 2004), implicit attitudes as measured by the IAT and related procedures have proved to be valid predictors of a wide range of behaviors that occur under suboptimal circumstances (e.g., Friese et al., 2009). They reach from the tendency to shoot at minorities under stress (Unkelbach et al., 2008) to overeating under the influence of alcohol (Hofmann and Friese, 2008). At times, implicit and explicit attitudes are positively related to one another, but under specified conditions, they

dissociate (Gawronski and Strack, 2004). It is exactly these conditions that allow deeper insights into the different operating mechanisms. Although attitudes are understood as predictors of behavior, the proposed duality has strongly directed the field's attention toward the dual representation of evaluations.

Duality in Behavior

Although many duality theories in social psychology focus on cognitive precursors of behavior such as attitude activation, person perception, stereotyping, and so on, at least [p. 103 ↓] some address the issue of the dual determination of behavior more directly. In fact, the idea that behavior can be determined by multiple factors which may sometimes be in conflict is at the heart of the implicit social cognition research program (e.g., Greenwald and Banaji, 1995; Wilson et al., 2000). Moreover, a great number of studies have addressed the dynamics of automatic social behavior and how it is controlled by environmental cues and inner regulatory mechanisms (e.g., Bargh, 2005; Bargh and Ferguson, 2000; Bargh et al., 1996). Conflicts between reflection and impulse have received most attention in the realm of research on self-control and impulsivity (e.g., Baumeister et al., 1998; Hofmann et al., 2009; Loewenstein, 1996; Metcalfe and Mischel, 1999).

In general, the concept of self-control implies that the execution of a behavior is preceded by a conflict between two forces. Economists were particularly concerned about conflicts between short-term and long-term consequences of decisions (Loewenstein and Elster, 1992) and in psychology, Walter Mischel and his colleagues (Metcalfe and Mischel, 1999; Mischel, 1974) have explored children's psychological dynamics in situations where they have the chance to increase their reward if they are able to wait. One of their determinants was the attention on the reward and the way it is mentally represented (Mischel et al., 1972). More recently, self-control has been identified as an activity that requires some unspecified type of energy (Baumeister et al., 1994). Using the analogy of a muscle, Baumeister and his colleagues suggest that self-control depletes this energy and replenishes it after some time. Although this notion has not been developed into a full-blown dual-process or system theory, it shares with them the assumption that there exist two types of activities that differ in the energy (or

cognitive capacity) they require. Specifically, there is energy consuming self-control and, by implication, the less taxing behavior that needs to be controlled.

An elaborate behavioral dual-system theory was proposed by Metcalfe and Mischel (1999) to account for the dynamics that undermine people's "willpower." Specifically, they postulate a "cool" system that operates on the basis of knowledge and is only indirectly related to behavior. The second, "hot," system is driven by emotions that serve as immediate determinants of impulsive behaviors. The model describes how the two systems interact and, more importantly, how they conflict when self-regulation breaks down. Thus, impulsive behavior is understood to be controlled not by the power of the will but by emotional forces that may sometimes be incompatible with willful intentions.

The Reflective-Impulsive Model

Most duality theories focus either on cognition (such as dual process models of persuasion or person perception) or on behavior (such as the strength model of self-regulation), but only very few models attempted to integrate cognitive and behavioral aspects (e.g., Epstein, 1990; Metcalfe and Mischel, 1999). Our own model (Strack and Deutsch, 2004) was developed with an integrative goal and particularly aims at specifying the interaction between the mechanisms of the systems that jointly determine behavior.

Building Blocks of the Model

The Reflective–Impulsive Model (RIM) describes the operating principles of two mental systems (reflective and impulsive) that jointly and interactively generate behavior. It is assumed that the two systems run in parallel and interact during the processing. The impulsive system is believed to be permanently engaged without consuming cognitive capacity; the reflective system is only active if enough cognitive capacity is available. Both systems, however, jointly determine how a perceptual input may eventually result in behavior.

The impulsive system The most fundamental link between perception and behavior was [p. 104 ↓] originally described as the ideo-motor principle (James, 1890; Lotze, 1852). It posits that without an intention or a goal, behavior may be elicited by merely perceiving or imagining certain content. As more recent findings suggest, activating concepts that include a motor component may exert similar effects (Rizzolatti et al., 2002). In the RIM, complex sensory-motor associations are an integral part of the impulsive system and are termed behavioral schemata. These structures resemble the concept of habits (Wood and Neal, 2007), and serve as the basis of overt behavior.

Associative structures. More frequently, the link between perception and behavior is less immediate but will be influenced by information that is already stored in memory like concepts and images. Stereotypes may exert such a mediating influence. For example, being exposed to a stereotype of an elderly person may cause people to reduce people's speed of walking (e.g., Bargh et al., 1996; Cesario et al., 2006). However, this influence seems to crucially depend on the actor's goal in a given situation (Cesario, et al., 2006). The absence of a pertinent decision or goal suggests that the behavior is elicited by a more complex structure in the impulsive system, which is assumed to emerge through past operations that connect perceptual input and behavioral output. Specifically, the impulsive system can be understood as an associative memory that closes the temporal gap between past and present. Based on the associative principles of frequency and recency, links between elements are assumed to emerge and are strengthened by their simultaneous activation. The general assumption of a passive creation of associative links, however, has recently been questioned (Mitchell et al., 2009). Instead, such links may mainly occur after more thorough encoding. There is, however, little doubt that, once associations have been formed, the activation of a particular element may be caused by another element with which it is strongly associated (like "hot" and "cold") or by several other elements that are jointly activated. In the previous example, having been exposed to the elderly stereotype may not have been sufficient to initiate the congruent walking behavior without inducing the person to get walking otherwise (Cesario et al., 2006).

Affect and valence. Beyond associative links to behavioral schemata, the RIM postulates a modifying role of valence and affect that is experienced in the course of the behavioral execution. First, the impulsive system is influenced by a mechanism that facilitates approach and avoidance. Specifically, it is subject to a "motivational

orientation" (Cacioppo et al., 1993; Neumann et al., 2003) that is triggered in a bidirectional fashion. On the one hand, the experience of positive or negative affect or the processing of positive or negative information may facilitate behavioral approach versus avoidance. In turn, these behaviors may facilitate the experience of the type of affect or the processing of the type of information that is compatible. In other words, decreasing the distance to a target (e.g., a person) will ease the experiencing of positive affect and the processing of positive information whereas increasing its distance will do the same for negative affect and information. It is important to note that on a spatial dimension, this may be achieved in two ways. Specifically, the distance may be changed by either moving away from the target (flight) or by causing the target to be moved away (fight).

It is important to note that changes in distance may be brought about by actual locomotion, by virtual operations on a computer screen (e.g., De Houwer et al., 2001; Markman and Brendl, 2005), or by symbolic actions like facial expressions and head movements (Neumann et al., 2005). Also, it is important to note that it has long been recognized that behaviors may influence evaluations. Specifically, Bem (1972) has posited that self-perception may be a causal determinant of attitudes. However, the proposed mechanism is quite different from the [p. 105 ↓] feedback of behavior on evaluations as outlined in the RIM. Specifically, self-perception theory proposes that people draw inferences from their behavior to the underlying attitudes. In contrast, the mechanisms of the RIM's impulsive system do not require inferences to be drawn. For example, it has been shown that the effects occurred even if experimental conditions prevented participants from identifying the meaning of their own behavior. The motivational orientation can therefore be seen as a general predisposition of the impulsive system that directs information processing and behavior by means of facilitation.

While approaching positive and avoiding negative targets has adaptive value, survival requires basic needs to be fulfilled. This suggests that the impulsive systems should also be responsive to specific needs such as food if the organism is in a state of deprivation. In fact, supporting evidence was found by Seibt et al. (2007) who found that the degree of approach motivation (as measured by lever movements) triggered by food stimuli was intensified in hungry compared to satiated participants. In summary,

affording automatic and fast adjustments to an environment is the speciality of the impulsive system.

The reflective system As much as speed is an important quality in reacting adaptively, fast behavior often prevents the actor from doing what is best. That is, the impulsive system has serious shortcomings that result from its fast and effortless processing. Specifically, it occurs with great rigidity and often fails to perform certain tasks. For example, the associative principle of frequency prevents single pieces of information from exerting an impact. Also, consequences that are temporarily remote, in both the future and the past, have less of an impact than they often should. In addition, the impulsive system has no place to allow for the transmission of knowledge without immediate exposure. That is, adaptation would be improved if people could learn from the experiences of others without the need to repeat them. This is particularly beneficial if such experiences have negative consequences.

The RIM assumes that primates are endowed with a second system that operates according to different principles to compensate for the shortcomings of the impulsive system. It is termed the “reflective system” and can be understood as being in charge of the generation and transformation of knowledge. The term “knowledge” is used in a relatively narrow sense in the RIM. More specifically, knowledge is defined as propositional statements about the world and the self, which contain truth-values. The cognitive processes of the reflective system assign and transform truth-values and generate new knowledge by inferences.

Returning to the previous example, being exposed to an old person, the impulsive system would automatically activate representations corresponding to the elderly stereotype. The reflective system, on the other hand, would form a proposition about the relation between this characteristic and the perceptual input, resulting in the knowledge that “this person is old.” Such knowledge can then be used for further transformations and inferences. Most simply, the truth-value can be reversed by applying the operation of negation. Thus, one might generate the subjective knowledge, or belief, that “this person is not old.” More important, the person may apply categorical knowledge about old people to draw inferences about characteristics of the target that are not accessible to immediate perception; for example, that “this person is hard of hearing.”

In other words, the reflective system generates an epistemic process that consists of several operations. It starts with linking a perceptual input to a name or a category (a process that we have referred to as “pointing and referring”; Strack and Deutsch, 2004). This results in a “propositional categorization” or “identification” which may then serve as a basis for a “noetic decision.” The nature of this decision may be factual and/or evaluative and lead to a “behavioral decision,” which is meant to reduce a [p. 106 ↓] discrepancy between a current state of the self and a possibility that has previously been positively evaluated. Given that decisions typically do not immediately result in behavior, it is plausible to assume specific processes to be responsible for mediating between decision and action. This process is labeled “intending” in the RIM (e.g., Gollwitzer, 1999). It links a behavioral decision to the actual execution of the behavior. The process of “intending” is particularly important to bridge the temporal gap between the “behavioral decision” and the execution of the behavior and to prevent the person from being preoccupied by reiterating or rehearsing the decision. Instead, it delegates the realization of the decision to the environment that reactivates it when the opportunity for the behavior is available. The process is concluded by the execution of the behavior or if the decision has become obsolete.

Interactions between the two systems Thus far, the two systems have been described separately, and we have identified both their central elements and their operational principles. In the following section, the interaction between the reflective and the impulsive system will be outlined by explaining how they relate at various stages of processing.

Most importantly, the operations in the reflective system need an informational basis. That is, to generate factual, evaluative, or behavioral decisions, it is necessary that relevant information is available. If the information comes from the environment, it needs to be categorized to serve as a basis for inferences. These concepts must have been stored at the time of categorization. Alternatively, the information does not come from concurrent environmental stimuli but from past experiences or from previous judgments. In all cases, information must have been stored to afford further processing. Different types of memory are assumed to be part of the two systems.

A working memory with a limited capacity and the capability to directly address its contents (Baddeley, 1986) is part of the reflective system, while the impulsive system

possesses a long-term store of unlimited capacity (e.g., Johnson and Hirst, 1991) that consists of associative structures.

The reflective system crucially depends on the associative store in that it provides the contents that are needed for its syllogistic operations. At a fundamental level, the execution of a propositional categorization depends on a category that must be available in memory. Even more so, inferences that are based on general knowledge can only be drawn if an appropriate schema is retrieved. This can be driven by the incoming stimulus if a category is strongly associated with the input. This category will be automatically activated and used for further processing in both the reflective and the impulsive system. In the absence of a strong link between the input and the category, the activation may be diluted and disappear. However, because the syllogistic operations in the reflective system rely on the input to be categorized, appropriate concepts need to be retrieved from the associative store in the impulsive system. In this case, the outcome depends not only on search cue but also on the activation potential (Johnson and Hirst, 1991) of the stored information, the accessibility of which relies on recency and frequency of its prior activation. Thus, each piece of information in the associative store has its own probability with which it enters into the operations of the reflective system. Thus, processes in the impulsive system exert a strong impact on reflective operations. Drawing from the seminal work by Higgins and his colleagues (Higgins, 1996), a person may be categorized as “reckless” only because people had previously been exposed to this concept, even if this had occurred in a different context.

From this logic, it follows that the causal impact may also go in the opposite direction. That is, the reflective use of a piece of information will alter its activation potential in the impulsive system. That is, thinking about a specific content will increase the probability that the same or related information will be [p. 107 ↓] retrieved at a later time. As trivial as this may sound, it has severe consequences for the generation of judgments. This becomes apparent in findings on anchoring effects in judgments under uncertainty (Higgins et al., 1977). In one of the experiments developed within this paradigm, people are required to generate two judgments in sequence. The first judgment is comparative in nature and uses the “anchor” as a standard of comparison. For example, judges may be asked to indicate if the Mississippi River is longer or shorter than 3,500 miles (high anchor) or 1,500 miles (low anchor). Subsequently, they have to estimate the river's actual length. Typically, this succession of two judgment tasks causes the

absolute judgment to be assimilated toward the standard of the preceding comparative judgment. In its original account, the effect was explained as the result of an “insufficient adjustment.” This explanation, however, begs many questions and fails to specify the underlying mechanisms (Tversky and Kahneman, 1974). From the vantage point of the RIM, the observed assimilation is the result of a selective accessibility of information at the time of the absolute judgment that is caused by the preceding comparative judgment. Specifically, we (Mussweiler et al., 1997) were able to demonstrate that to generate comparative judgments, people act as if the standard of comparison is a single hypothesis (Klayman and Ha, 1987) and seek confirmatory information. Even if this biased search results in a rejection of the hypothesis, and in a response that is qualified by the standard, it also leads to a selective accessibility of information at the time of the absolute judgment. As a consequence, the information that is then retrieved is biased and the resulting judgment is assimilated toward the standard of the comparative task. Assuming such an interaction between the two systems allows for an understanding of judgmental biases that is linked to more general psychological mechanisms that are responsible for a great variety of phenomena. This parsimony is an important achievement of dual-system models that are sufficiently specified in its processes.

Two types of interaction between the reflective and the impulsive system are described by these mechanisms. On the one hand, the existing activation potential determines the likelihood with which a content is used for reflective operations; at the same time, reflective processes alter the accessibility of information for operations of the impulsive system.

Synergistic and antagonistic operations It is an important feature of the RIM that the final pathway to behavior is jointly used by both systems. In the reflective system, behavior is the result of a noetic decision about the desirability of an outcome and the feasibility with which it can be achieved (Ajzen, 1991; Mussweiler and Strack, 1999). As has been outlined previously, the information on which the decision is based may be influenced by the selective accessibility of information in the impulsive system. However, the translation of a behavioral decision into overt behavior may not occur in an immediate fashion. Instead, the impulsive system may activate behavioral schemata that are incompatible with the behavioral decision and thus interfere with its execution.

This suggests that the compatibility of the two systems determines the strength of a particular behavior. Specifically, if the activities in both systems lead to the activation of the same schema, it will be easier to execute the behavior. It may be further facilitated if the impulsive system creates fluency and a feeling of positive affect (Bandura, 1977; Topolinski and Strack, 2009)

In contrast, if the two systems activated incompatible schemata or if the reflective system inhibits impulsive behaviors, the two systems compete with one another. At the same time, feelings of temptation and conflict may arise. Which system will prevail depends on situational and dispositional characteristics. In the following, we shall describe such determinants along with pertinent empirical evidence.

[p. 108 ↓]

Determinants of Impulsive versus Reflective Behaviors

Situational determinants For the reflective system to operate, cognitive capacity is necessary. Thus, being preoccupied by another mental activity will decrease the probability that behaviors are controlled by reflective operations. Specifying this distracting influence, Baumeister and colleagues (e.g., Winkielman and Cacioppo, 2001) have invoked the metaphor of energy that exists as a limited resource and will be depleted by a mentally or physically taxing task. Also, they claim that it takes time until the energy is replenished. This assumption is largely consistent with the RIM, which provides a conceptual basis to account for mechanisms of self-control (Baumeister and Heatherton, 1996).

Combining Baumeister' energy model with the RIM, Hofmann et al. (2007) have depleted participants' cognitive resources by requiring them to suppress a negative emotion that had been induced by a movie. Subsequently, they took part in a taste test in which M&M candies had to be evaluated in several dimensions. However, the researchers' interest was not in participants' evaluations but in their eating behavior. Specifically, they wanted to find out if they were more likely to act against their

dieting intentions under depletion. This was the case. Participants' eating behavior corresponded to their dietary restraint standards much more if they had not previously been depleted. If they had, however, eating behavior was related to their automatic affective reactions to M&Ms that was previously assessed by a variant of the IAT (Hofmann et al., 2007). This finding was replicated for other types of food and for alcohol consumption (Greenwald et al., 1998). In summary, these results suggest that behavior is controlled by reflection only if the necessary cognitive resources are available. Otherwise, it is more likely to be elicited by affectively loaded impulses which can be assessed by implicit measures capturing the affect that is associatively linked to the target of the behavior.

In a more direct way, Friese et al. (2008) depleted participants' cognitive capacity by giving and having them execute a secondary memory task. Specifically, while they were invited to choose between chocolate bars and fruits, they had to memorize either an eight-digit number (high cognitive load) or a one-digit number (low cognitive load). As expected, people's actual behavior corresponded more strongly with their explicit evaluation if the secondary task was easy than if it was difficult. Again, under the difficult, high-load condition, automatic affective reactions (measured by the IAT) were more likely to predict the choice.

Finally, the behavioral conflict between reflective and impulsive determination can be resolved by alcohol consumption. Specifically, impulsive influences are assumed to be stronger under alcoholic intoxication. This prediction was tested by Hofmann and Friese (2008) in a study that again employed a taste test in which participants had to evaluate different kinds of candy products. Unlike in the previous study, there was a preceding "product test" that involved orange juice that was either blended with vodka or not. Again, the amount of candy that was consumed was the dependent variable. The results conceptually replicate those of the experiment by Hofmann et al. (2007). That is, alcohol consumption was equivalent with the depletion (i.e., affect suppression) manipulation in that participants' dietary restraint standards were less and associative affect was more effective in determining eating behavior.

Dispositional determinants Of course, a considerable part of the variation in people's tendency toward one of the two conflicting behaviors is determined by personal characteristics that may have emerged as a function of inheritance or learning. For

example, such interpersonal differences can be found for the domain of cognitive capacity. Specifically, working memory capacity has been assessed and studied in its relation to behaviors that often occur in an impulsive fashion. As a [p. 109 ↓] consequence, people's actual behavior can be better predicted using explicit evaluative assessments if their cognitive capacity is high. If it is low, impulsive evaluative reactions should be better predictors.

This pattern was tested using erotic stimuli. The authors (Hofmann and Friese, 2008) recorded how long participants would look at erotic versus art pictures. In addition, their working memory capacity was assessed. The findings demonstrate that for those people who had a high memory capacity attention toward the erotic pictures was best predicted by explicit evaluations. In contrast, the IAT was the best predictor for those whose working memory capacity had been low.

Applicability to Social Issues

Beyond providing a conceptual structure to integrate diverse psychological phenomena, the RIM is relevant to various issues in the real world. This is particularly apparent in the reception that the model received in several disciplines of applied psychology.

One outstanding example is the field of health psychology where interventions play an important role in preventing people from engaging in harmful behaviors. This has been particularly evident in the case of addiction. For example, the RIM implies that addictive behaviors are predominantly controlled by the impulsive system which does not have the ability to negate propositional contents (e.g., Deutsch et al., 2006; Gilbert, 1991). Instead, the affirmative content will be activated. As a consequence, persuasive communications with demands that include a negation (e.g., "Don't smoke!") should be less effective than communications that do not mention the harmful behavior at all. A study comparing media campaigns against smoking suggests that this is, in fact, the case (Farrelly et al., 2002). As a recommendation, the RIM suggests that in the case of impulsive behaviors, just saying no is not enough (Gawronski et al., 2008).

Implications of the model have been applied to the harmful consumption of alcoholic beverages and addictive behaviors in general (Deutsch and Strack, 2006b). Reinout

Wiers and his colleagues have found that alcoholic beverages affected drinkers' motivational orientation. For these persons, behaviors approaching alcoholic beverages were found to be facilitated (Wiers et al., 2009). More importantly, counterimpulsive training proved effective even over an extended period of time. Specifically, heavy drinkers' consumption of alcohol was reduced if they had to engage in avoidance behaviors (pushing a lever away) while being exposed to pictures of alcoholic beverages (Wiers et al., 2010).

In the health domain, the conflict between dietary restraint and the impulse to eat caloric food is particularly obvious. The choice between an apple and a chocolate ice cream is not the same as between an apple and a pear, because impulses are more likely to be involved in the first case. Wilhelm Hofmann and his colleagues (for a review, see Hofmann et al., 2009) have convincingly demonstrated that reflective control of tempting behaviors breaks down if the psychological operating conditions of reflective control are undermined or not fulfilled; for example, when people are temporarily depleted of self-regulatory resources, under the influence of alcohol, or low in chronic working memory capacity. Thus, the colloquial "strength of willpower" in successfully resisting temptation appears to be grounded in specific antecedents of reflective operations related to concepts of central executive functioning. In cases where these functions are temporarily or chronically ineffective, impulsive processes appear to gain the upper hand in determining behavior (e.g., Allan et al., 2010; Hofmann et al., 2008; Hoefling and Strack, 2008).

The term "impulsivity" has also been used to describe behaviors in the domain of marketing. There, "impulse buying" (Vohs and Faber, 2007) is understood as a type of economic exchange that is not determined by [p. 110 ↓] rational anticipation of the outcome of the action but by an immediate urge to appropriate a product. Of course, the RIM is particularly capable of explaining the interaction between rational choice and impulsive behavior which is at the core of economic phenomena (Strack et al., 2006).

More generally, the RIM seeks to integrate previous research and theorizing on stereotypes and attitudes, and to connect these constructs to the general assumptions about judgment, motivation, and behavior. We hope that this integration may be helpful in improving and stimulating research aimed at reducing intergroup conflict and discrimination.

Conclusion

The reported findings show that impulsive and reflective types of behavior are systematically related to determinants that can be situational in nature or may reside as a stable characteristic in the person. In both cases, the results strongly suggest that human behavior is not a unified phenomenon but is determined by different psychological mechanisms. They are assumed to belong to different systems whose operation depends on specific conditions.

The structure of the RIM allows studying the synergistic or antagonistic influences of the two systems and to explore the conditions of their operation. Some of these have been described in this chapter. Others are open to future explorations. For example, it might be possible to facilitate reflective versus impulsive processing using appropriate induction tasks. Specifically, this might be achieved by inducing concrete versus abstract representations that have been found to induce feelings versus evaluations (Strack et al., 1985). Recent research (Fujita and Han, 2009) suggests that this may be generalized across contents. As a consequence, one might predict that impulsive behaviors are caused by concrete representations whereas their reflective regulations may be more successful if abstract thinking is prevalent.

Other programs of research may focus on the relationships between evaluations and goals (Markel, 2009), evaluative conditioning versus inferences (Corneille et al., 2009; Hofmann et al., 2008), frustration and behavioral orientation (Krieglmeyer et al., submitted) or the role of counterimpulse training in the domain of addiction (Mitchell, et al., 2009).

In summary, it seems as if with new methodologies and theoretical models, the old theme of the dual determination of human behavior has become a central topic of psychological research, which has important basic and applied implications. This allows the achievement of a deeper understanding of those components that go beyond the anticipation and evaluation of outcomes and includes unplanned behaviors. In particular, it reconsiders impulsive elements that have long been neglected in social psychology. More importantly, it integrates them into a coherent model of human behavior in which the contribution of different components is explained as a function

of situational and dispositional antecedents. As a consequence, human behavior is understood as a blend of constituents that is rarely “process pure.” Instead, planned and spontaneous components are part and parcel of any behavioral manifestation. To further disentangle their contributions and explain their operation under specified circumstances is a challenge for future endeavors.

Appendix: Personal Histories of the Rim

Reporting one's own intellectual history is not without dangers because it lends itself to various biases. Admitting one's own thinking as autobiographical evidence presupposes that our memory for our own cognitive activity is solely determined by what has actually happened. Unfortunately, this is not the case [p. 111 ↓] and reconstructive endeavors (e.g., Strack and Bless, 1994) may greatly contribute to the recollection and distort the validity of the reminiscence for the benefit of a coherent narrative. However, these distortions tend to remain largely unnoticed by the protagonists and may require corrections from critical observers. Given these caveats, here is how we came to develop the Reflective-Impulsive Model.

F.S.: Quite some time ago, I was excited about the controversial discussion on the possibility of mental imagery (e.g., Paivio, 1969). Following this exchange, I was convinced that there exists something like a visual representation simply because people who try to visually imagine something often cover their eyes but never their ears. This suggested to me that relation between visual representation and visual perception is not metaphorical but psychologically substantive. That is, to “see” somebody with one's mind's eye is qualitatively similar to the result of the perceptual act of seeing. Luckily, this view was subsequently supported in a review article by Martha Farah (1988) in which evidence was collected demonstrating that similar cortical areas are involved in visual imagery and visual perception.

Thus, it seemed reasonable to assume that perceptual representations do not completely vanish once they have left the sensory register. Their quality may continue to exert its effect if people actively imagine a target or event. At the same time,

categorical or symbolic representation may coexist and the world may be represented in two ways: as categorical knowledge or as a perceptual experience. Of course, in reality both types may coexist and supplement one another. Still, this distinction convinced me about the possibility that the world may be internally represented in different ways.

In some early studies in collaboration with Norbert Schwarz (Strack et al., 1985), I tried to manipulate the type of representation by inducing participants to think about past life events in either a concrete ("How did it happen?") or in an abstract ("Why did it occur?") fashion. We found that in the concrete condition, participants experienced more affect that was congruent with the valence of the event than in the abstract condition. These findings suggest that perceptual experiences may be similar to immediate experiences that are elicited by imagery.

The two types of representations may sometimes be in conflict. This is apparent in visual illusions, like the Müller-Lyer Illusion. Here, we perceive the segment of the line between the open arrows to be shorter than the segment between the pointed arrows. At the same time our perceptual experience may be at conflict with our knowledge that the two lines are of equal length. Because they have divergent implications, we are aware of the dual representation of the target.

In a subsequent book chapter (Strack, 1992), I have made a case for a dual representation of attitudes. In particular, I have argued that formation of attitudes involves both an experiential and an informational route and that the inclusion of experiences will be subject to correction if they are not "representative" for the judgment.

Yet another type of duality fed into the theory that is described in the chapter. It comes from research on metacognition I conducted in collaboration with Herbert Bless (Strack and Bless, 1994) and Jens Förster (Strack and Förster, 1998). Here, it was the interaction between recollective experience and metacognitive knowledge that was shown to jointly contribute to judgments of previous occurrence.

To account for these observations in a systematic way, it has long been my plan to integrate these various insights into a model that focuses on the interaction of different representations and concomitant operating mechanisms in information processing.

Moreover I wanted to link affective and cognitive processes with behavior as the ultimate outcome.

However, this ambitious goal could only be reached with a highly competent and motivated collaborator. Roland Deutsch, whom I knew since he was an undergraduate [p. 112 ↓] student, turned out to be the ideal person with whom to attempt the project. Luckily, I was able to persuade him to come aboard. This resulted in numerous discussions and many revisions of a manuscript that was finally published in 2004. Independent of how the end product is eventually evaluated, the joint way to get there was one of the most stimulating and challenging phases of my entire career.

R.D.: While working towards my degree in psychology, I chiefly focused on theories on animal and human behavior that were developed in the field of experimental psychology. Back then (and still), I was fascinated by connectionist models of behavioral performance, and by associative learning more broadly defined. At the same time, I delved into the rich literature on the regularities of social judgment and decision making.

Although both fields attracted me equally, I had the impression that theory and research in these fields were conducted largely in an independent manner – a state of affairs that I deemed highly undesirable. As a consequence, theories that apparently seemed to make a step toward reconciling the realms of judgment and decision making on one hand and basic learning mechanisms on the other hand strongly attracted my attention. One of these theories was the theory of a supervisory attentional system by Norman and Shallice (1986), which served as a working model for my own psychological thinking for quite a while.

After I decided to pursue my doctoral studies in the field of social psychology, the idea of integrating judgmental and behavioral theories was still a strong motive. And although social psychology already had a lot of “duality theories” to offer, most of these theories focused on different types of social judgments without thoroughly trying to integrate basic processes of motivation and behavior control.

Primed by this history of studies and interests, I entered a lab meeting in the fall of 1999, which I still remember very vividly. I was still relatively new to the department, and

I had experienced Fritz Strack mainly as a person who contributed to the lab meeting as supervisor and discussant. But for this special occasion, he had invited all members to attend an introductory talk to his plans on developing an integrative theory of social judgment and social behavior.

Much of these ideas seemed to represent what I had been looking for in the past. At the same time, Fritz pointed out that substantial work was still necessary to turn the ideas conveyed in this talk into an elaborate theory, and he invited the audience to join his efforts. I immediately knew that I wanted to be part of this project. The resulting hours of stimulating and challenging discussions, massive literature searches, and many fine-grained revisions of what was to become the final Reflective-Impulsive Model where a formative experience, which was and still is the basis for my academic work

Note

1 Another example for a cultural manifestation of the duality of human behavior is a passage from Goethe's *Faust* in which the protagonist complains that "two souls, alas!" reside within his "breast and each withdraws from and repels its brother" (*Faust I*, 2, 307). Here, the two principles are given equal status and the dissociation between them is depicted as a regrettable experience.

References

Ainslie, G. (1975) Specious reward: A behavioral theory of impulsiveness and impulse control . Psychological Bulletin , 82, 463. <http://dx.doi.org/10.1037/h0076860>

Ainslie, G. (2005) Precis of breakdown of will . Behavioral and Brain Sciences , 28, 635–673.

Ajzen, I. (1991) The theory of planned behavior . Organizational Behavior and Human Decision Processes , 50, 179–211. <http://dx.doi.org/10.1016/0749-5978%2891%2990020-T>

Allan, J.L., Johnston, M. and Campbell, N. (2010) Unintentional eating. What determines goal-congruent chocolate consumption . *Appetite* , 54, 422–425. <http://dx.doi.org/10.1016/j.appet.2010.01.009>

Asendorpf, J.B. and van Aken, M.A. (1999) Resilient, overcontrolled, and undercontrolled personality prototypes in childhood: Replicability, predictive power, and the trait-type issue . *Journal of Personality and Social Psychology* , 77, 815–832. <http://dx.doi.org/10.1037/0022-3514.77.4.815>

Baddeley, A.D. (1986) Working Memory Oxford: Clarendon Press.

Bandura, A. (1977) Social Learning Theory Oxford: Prentice-Hall.

Bargh, J.A. (1994. The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition . In R.S. Wyer, ed. and T.K. Srull (eds), *Handbook of Social Cognition* , 1, 1–40. Hillsdale, NJ: Erlbaum.

Bargh, J.A. (2005) Bypassing the will: Toward demystifying the nonconscious control of social behavior . In R.R. Hassin, ed. , J.S. Uleman, ed. and J.A. Bargh (eds), *The New Unconscious* , pp. 37–58. Oxford: Oxford University Press.

Bargh, J.A., Chen, M. and Burrows, L. (1996) Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action . *Journal of Personality and Social Psychology* , 71, 230–244. <http://dx.doi.org/10.1037/0022-3514.71.2.230>

Bargh, J.A. and Ferguson, M.J. (2000) Beyond behaviorism: On the automaticity of everyday life . *Psychological Bulletin* , 126, 925–945. <http://dx.doi.org/10.1037/0033-2909.126.6.925>

Baumeister, R.F., Bratslavsky, M., Muraven, M. and Tice, D.M. (1998) Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology* , 74, 1252–1265. <http://dx.doi.org/10.1037/0022-3514.74.5.1252>

Baumeister, R.F. and Heatherton, T.F. (1996) Self-regulation failure: An overview . *Psychological Inquiry* , 7, 1–15. http://dx.doi.org/10.1207/s15327965pli0701_1

Baumeister, R.F., Heatherton, T.F. and Tice, D.M. (1994) *Losing Control: How and Why People Fail at Self-regulation*. San Diego: Academic Press.

Bem, D.J. (1972) Self-perception theory . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 6, 1–62. New York: Academic Press.

Blanton, H. and Jaccard, J. (2006) Arbitrary metrics in psychology . *American Psychologist* , 61, 27–41. <http://dx.doi.org/10.1037/0003-066X.61.1.27>

Block, J.H. and Block, J. (1980) The role of ego-control and ego-resiliency in the organization of behavior . In W.A. Collins (ed.), *Development of Cognition, Affect, and Social Relations* , 13, 39–101. Hillsdale, NJ: Lawrence Erlbaum.

Cacioppo, J.T., Priester, J.R. and Berntson, G.G. (1993) Rudimentary determinants of attitudes. II: Arm flexion and extension have differential effects on attitudes . *Journal of Personality and Social Psychology* , 65, 5–17. <http://dx.doi.org/10.1037/0022-3514.65.1.5>

Carver, C.S. (2005) Impulse and constraint: Perspectives from personality psychology, convergence with theory in other areas, and potential for integration . *Personality and Social Psychology Review* , 9, 312–333. http://dx.doi.org/10.1207/s15327957pspr0904_2

Cesario, J., Plaks, J.E. and Higgins, E.T. (2006) Automatic social behavior as motivated preparation to interact . *Journal of Personality and Social Psychology* , 90, 893–910. <http://dx.doi.org/10.1037/0022-3514.90.6.893>

Chaiken, S. (1980) Heuristic versus systematic information processing and the use of source versus message cues in persuasion . *Journal of Personality and Social Psychology* , 39, 752–766. <http://dx.doi.org/10.1037/0022-3514.39.5.752>

Chaiken, S., ed. and Trope, Y. (eds) (1999). *Dual-process Theories in Social Psychology* . New York: Guilford.

Chen, M. and Bargh, J. (1999) Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus . *Personality and Social Psychology Bulletin* , 25, 215–224. <http://dx.doi.org/10.1177/0146167299025002007>

Corneille, O., Yzerbyt, V.Y., Pleyers, G. and Mussweiler, T. (2009) Beyond awareness and resources: Evaluative conditioning may be sensitive to processing goals . *Journal of Experimental Social Psychology* , 45, 279–282. <http://dx.doi.org/10.1016/j.jesp.2008.08.020>

Costa, P.T., Jr. and McCrae, R.R. (2009) The five-factor model and the NEO inventories . In J.N. Butcher (ed.), *Oxford Handbook of Personality Assessment* , pp. 299–322. New York: Oxford University Press.

De Houwer, J., Crombez, G., Baeyens, F. and Hermans, D. (2001) On the generality of the affective Simon effect . *Cognition and Emotion* , 15, 189–206.

De Houwer, J. and Moors, A. (2007). How to define and examine the implicitness of implicit measures. In B. Wittenbrink, ed. and N. Schwarz (eds), *Implicit Measures of Attitudes* , pp. 179–194. New York: Guilford Press.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A. and Moors, A. (2009) Implicit measures: A normative analysis and review . *Psychological Bulletin* , 135, 347–368. <http://dx.doi.org/10.1037/a0014211>

Deutsch, R., Gawronski, B. and Strack, F. (2006) At the boundaries of automaticity: Negation as reflective operation . *Journal of Personality and Social Psychology* , 91, 385–405. <http://dx.doi.org/10.1037/0022-3514.91.3.385>

Deutsch, R. and Strack, F. (2006a) Duality models in social psychology: From dual processes to interacting systems . *Psychological Inquiry* , 17, 166–172. http://dx.doi.org/10.1207/s15327965pli1703_2

Deutsch, R. and Strack, F. (2006b) Reflective and impulsive determinants of addictive behavior . In R.W. Wiers, ed. and A.W. Stacy (eds), *Handbook of Implicit Cognition and Addiction* . Thousand Oaks, CA: Sage. <http://dx.doi.org/10.4135/9781412976237>

Devine, P.G. (1989) Automatic and controlled processes in prejudice: The role of stereotypes and personal beliefs . In A.R. Pratkanis, ed. , S.J. Breckler, ed. and A.G. Greenwald (eds), *Attitude Structure and Function* , pp. 181–212. Hillsdale, NJ: Lawrence Erlbaum Associates.

Dimberg, U., Thunberg, M. and Grunedal, S. (2002) Facial reactions to emotional stimuli: Automatically controlled emotional responses . *Cognition and Emotion* , 16, 449–472. <http://dx.doi.org/10.1080/02699930143000356>

Dunton, B.C. and Fazio, R.H. (1997) An individual difference measure of motivation to control prejudiced reactions . *Personality and Social Psychology Bulletin* , 23, 316–326. <http://dx.doi.org/10.1177/0146167297233009>

Epstein, S. (1990) Cognitive-experiential self-theory . In L.A. Pervin (ed.), *Handbook of Personality. Theory and Research* , 165–192. New York: Guilford.

Farah, M.J. (1988) Is visual imagery really visual? Overlooked evidence from neuropsychology . *Psychological Review* , 95, 307–317. <http://dx.doi.org/10.1037/0033-295X.95.3.307>

Farrelly, M.C., Healton, C.G., Davis, K.C., Messeri, P., Hersey, J.C. and Haviland, M.L. (2002) Getting to the truth: Evaluating national tobacco countermarketing campaigns . *American Journal of Public Health* , 92, 901–907. <http://dx.doi.org/10.2105/AJPH.92.6.901>

Fazio, R.H., Sanbonmatsu, D.M., Powell, M.C. and Kardes, F.R. (1986) On the automatic activation of attitudes . *Journal of Personality and Social Psychology* , 50, 229–238. <http://dx.doi.org/10.1037/0022-3514.50.2.229>

Fiedler, K., Messner, C. and Bluemke, M. (2006) Unresolved problems with the 'I', the 'A', and the 'T': A logical and psychometric critique of the Implicit Association Test (IAT) . In W.H.M. Stroebe (ed.), *European Review of Social Psychology* , 17, 74–147. New York: Psychology Press.

Frederick, S., Loewenstein, G.F. and O'Donoghue, T. (2003) Time discounting and time preference: A critical review . In G.-F. Loewenstein, ed. , D. Read, ed. , and R.

Baumeister (eds), Time and Decision: Economic and Psychological Perspectives on intertemporal Choice , pp. 13–86. New York: Russell Sage Foundation.

Freud, S. (1927) The Ego and the Id . London: Hogarth.

Freud, S. (1993) Massenpsychologie und Ich-Analyse . Frankfurt: Fischer Taschenbuch Verlag. (Originally published in 1921.)

Friese, M., Hofmann, W. and Schmitt, M. (2009) When and why do implicit reaction time measures predict behavior? Empirical evidence for the moderating role of motivation, opportunity, and process reliance . European Review of Social Psychology , 19, 285–338. <http://dx.doi.org/10.1080/10463280802556958>

Friese, M., Hofmann, W. and Wänke, M. (2008) When impulses take over: Moderated predictive validity of explicit and implicit attitude measures in predicting food choice and consumption behaviour . British Journal of Social Psychology , 47, 397–419. <http://dx.doi.org/10.1348/014466607X241540>

Fujita, K. and Han, H.A. (2009) Moving beyond deliberative control of impulses: The effect of construal levels on evaluative associations in self-control conflicts . Psychological Science , 20, 799–804. <http://dx.doi.org/10.1111/j.1467-9280.2009.02372.x>

Funder, D.C. and Block, J. (1989) The role of ego-control, ego-resiliency, and IQ in delay of gratification in adolescence . Journal of Personality and Social Psychology , 57, 1041–1050. <http://dx.doi.org/10.1037/0022-3514.57.6.1041>

Gawronski, B., Deutsch, R., Mbirkou, S., Seibt, B. and Strack, F. (2008) When ‘just say no’ is not enough: Affirmation versus negation training and the reduction of automatic stereotype activation . Journal of Experimental Social Psychology , 44, 370–377. <http://dx.doi.org/10.1016/j.jesp.2006.12.004>

Gawronski, B. and Strack, F. (2004) On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes . Journal of Experimental Social Psychology , 40, 535–542. <http://dx.doi.org/10.1016/j.jesp.2003.10.005>

Gilbert, D.T. (1991) How mental systems believe . American Psychologist , 46, 107–119. <http://dx.doi.org/10.1037/0003-066X.46.2.107>

Gollwitzer, P.M. (1999) Implementation intentions: Strong effects of simple plans . American Psychologist , 54, 493–503. <http://dx.doi.org/10.1037/0003-066X.54.7.493>

Greenwald, A.G. and Banaji, M.R. (1995) Implicit social cognition: Attitudes, self-esteem, and stereotypes . Psychological Review , 102, 4. <http://dx.doi.org/10.1037/0033-295X.102.1.4>

Greenwald, A.G., McGhee, D.E. and Schwartz, J.L.K. (1998) Measuring individual differences in implicit cognition: The Implicit Association Test . Journal of Personality and Social Psychology , 74, 1464–1480. <http://dx.doi.org/10.1037/0022-3514.74.6.1464>

Greifeneder, R. and Bless, H. (2007) Relying on accessible content versus accessibility experiences: The case of processing capacity . Social Cognition , 25, 853–881. <http://dx.doi.org/10.1521/soco.2007.25.6.853>

Higgins, E.T. (1996) Knowledge activation: Accessibility, applicability, and salience . In E.T. Higgins, ed. and A.W. Kruglanski (eds), Social Psychology. Handbook of Basic Principles , pp. 133–168. New York: Guilford Press.

Higgins, E.T., Rholes, W.S. and Jones, C.R. (1977) Category accessibility and impression formation . Journal of Experimental Social Psychology , 13, 141–154. <http://dx.doi.org/10.1016/S0022-1031%2877%2980007-3>

Hoefling, A. and Strack, F. (2008) The tempting effect of forbidden foods: High calorie content evokes conflicting implicit and explicit evaluations in restrained eaters . Appetite , 51, 681–689. <http://dx.doi.org/10.1016/j.appet.2008.06.004>

Hofmann, W. and Friese, M. (2008) Impulses got the better of me: Alcohol moderates the influence of implicit attitudes toward food cues on eating behavior . Journal of Abnormal Psychology , 117, 420–427. <http://dx.doi.org/10.1037/0021-843X.117.2.420>

Hofmann, W., Friese, M. and Strack, F. (2009) Impulse and self-control from a dual-systems perspective . *Perspectives on Psychological Science* , 4, 162–176. <http://dx.doi.org/10.1111/j.1745-6924.2009.01116.x>

Hofmann, W., Gschwendner, T., Friese, M., Wiers, R.W. and Schmitt, M. (2008) Working memory capacity and self-regulation: Toward an individual differences perspective on behavior determination by automatic versus controlled processes . *Journal of Personality and Social Psychology* , 95, 962–977. <http://dx.doi.org/10.1037/a0012705>

Hofmann, W., Rauch, W. and Gawronski, B. (2007) And deplete us not into temptation: Automatic attitudes, dietary restraint, and self-regulatory resources as determinants of eating behavior . *Journal of Experimental Social Psychology* , 43, 497–504. <http://dx.doi.org/10.1016/j.jesp.2006.05.004>

James, W. (1890) *The Principles of Psychology* (Vol. 1) . New York: Dover. <http://dx.doi.org/10.1037/10538-000>

Johnson, M.K. and Hirst, W. (1991) Processing subsystems of memory . In R.G.W.H.J. Lister (ed.), *Perspectives on Cognitive Neuroscience* , pp. 197–217. New York: Oxford University Press.

Kahneman, D. and Tversky, A. (1982) The simulation heuristic . In D. Kahneman, ed. , P. Slovic, ed. and A. Tversky (eds), *Judgment Under Uncertainty. Heuristics and Biases* , pp. 201–208. Cambridge: Cambridge University Press.

Klayman, J. and Ha, Y. (1987) Confirmation, disconfirmation, and information in hypothesis testing . *Psychological Review* , 94, 211–228. <http://dx.doi.org/10.1037/0033-295X.94.2.211>

Koriat, A. and Levy-Sadot, R. (1999) Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one's own knowledge . In S. Chaiken, ed. and Y. Trope (eds), *Dual-process Theories in Social Psychology* , pp. 483–502. New York: Guilford Press.

Kosslyn, S.M. (1980) *Image and Mind* . Cambridge, MA: Harvard University Press.

Krieglmeyer, R., Deutsch, R. and Strack, F. (submitted) The effect of frustration on behavioral tendencies of approach and avoidance.

Kruglanski, A.W. and Thompson, E.P. (1999) Persuasion by a single route: A view from the unimodel . *Psychological Inquiry* , 10, 83–109. <http://dx.doi.org/10.1207/S15327965PL100201>

Le Bon, G. (1895) *The Crowd* New York: Viking Press.

Loewenstein, G.F. (1996) Out of control: Visceral influences on behavior . *Organizational Behavior and Human Decision Processes* , 65, 272. <http://dx.doi.org/10.1006/obhd.1996.0028>

Loewenstein, G.F. and Elster, J. (1992) *Choice Over Time* . New York, NY: Russell Sage Foundation.

Loewenstein, G. and Thaler, R.H. (1997) Intertemporal choice . In W.M. Goldstein, ed. and R.M. Hogarth (eds), *Research on Judgment and Decision Making: Currents, Connections, and Controversies* , pp. 365–378. New York: Cambridge University Press.

Loewenstein, G. and Thaler, R.H. (2005) Anomalies: Intertemporal choice . In M. Bazerman (ed.) *Negotiation, Decision Making and Conflict Nanagement* , pp. 428–440. Northampton, MA: Edward Elgar Publishing.

Lotze, H. (1852) *Medizinische Psychologie oder die Physiologie der Seele* . Leipzig: Weidmann'sche Buchhandlung.

Markel, P. (2009) If it's to be, it starts with me! The bidirectional relation between goals and the self . PhD dissertation, University of Würzburg.

Markman, A.B. and Brendl, C.M. (2005) Constraining theories of embodied cognition . *Psychological Science* , 16, 6–10. <http://dx.doi.org/10.1111/j.0956-7976.2005.00772.x>

McClure, S.M., Laibson, D.I., Loewenstein, G. and Cohen, J.D. (2004) Separate neural systems value immediate and delayed monetary rewards . *Science* , 306, 503–507. <http://dx.doi.org/10.1126/science.1100907>

Metcalfe, J. and Mischel, W. (1999) A hot/cool system analysis of delay of gratification: Dynamics of willpower . *Psychological Review* , 106, 3–19. <http://dx.doi.org/10.1037/0033-295X.106.1.3>

Mischel, W. (1974) Processes in delay of gratification . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 7, 249–292. San Diego: Academic.

Mischel, W., Ebbesen, E.B. and Zeiss, A.R. (1972) Cognitive and attentional mechanisms in delay of gratification . *Journal of Personality and Social Psychology* , 21, 204–218. <http://dx.doi.org/10.1037/h0032198>

Mitchell, C.J., De Houwer, J. and Lovibond, P.F. (2009) The propositional nature of human associative learning . *Behavioral and Brain Sciences* , 32, 183–198. <http://dx.doi.org/10.1017/S0140525X09000855>

Mussweiler, T., Förster, J. and Strack, F. (1997) Anchoring effects and the applicability of anchor-consistent information: A selective accessibility model . *Zeitschrift für Experimentelle Psychologie* , 44, 589–615.

Mussweiler, T. and Strack, F. (1999) Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model . *Journal of Experimental Social Psychology* , 35, 136–164. <http://dx.doi.org/10.1006/jesp.1998.1364>

Neumann, R., Förster, J. and Strack, F. (2003) Motor compatibility: The bidirectional link between behavior and evaluation . In J. Musch, ed. and K.C. Klauer (eds), *The Psychology of Evaluation* , pp. 371–391. London: Lawrence Erlbaum Associates.

Neumann, R., Hess, M., Schulz, S.M. and Alpers, G.W. (2005) Automatic behavioural responses to valence: Evidence that facial action is facilitated by evaluative processing . *Cognition and Emotion* , 19, 499–513.

Norman, D.A. and Shallice, T. (1986) Attention to action. Willed and automatic control of behavior . In R.J. Davidson, ed. , G.E. Schwartz, ed. and D. Shapiro (eds), *Consciousness and Self Regulation: Advances in Research* , pp. 1–18. New York: Plenum.

Paivio, A. (1969) Mental imagery in associative learning and memory . Psychological Review , 76, 241–263. <http://dx.doi.org/10.1037/h0027272>

Petty, R.E. and Cacioppo, J.T. (1986) The elaboration likelihood model of persuasion . In L. Berkowitz (ed.) Advances in Experimental Social Psychology , 19, 123–205. San Diego: Academic Press.

Rizzolatti, G., Craighero, L. and Fadiga, L. (2002) The mirror system in humans . In M.I.G.V. Stamenov (ed.), Mirror Neurons and the Evolution of Brain and Language , pp. 37–59. Amsterdam: John Benjamins.

Rothermund, K. and Wentura, D. (2004) Underlying processes in the Implicit Association Test (IAT): Dissociating salience from associations . Journal of Experimental Psychology. General , 133, 139–165. <http://dx.doi.org/10.1037/0096-3445.133.2.139>

Rowe, C.J. (1998) Plato: Phaedrus, with Translation and Commentary . Warminster: Aris and Philips.

Schneider, W. and Shiffrin, R.M. (1977) Controlled and automatic human information processing: I. Detection, search, and attention . Psychological Review , 84, 1–66. <http://dx.doi.org/10.1037/0033-295X.84.1.1>

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H. and Simons, A. (1991) Ease of retrieval as information: Another look at the availability heuristic . Journal of Personality and Social Psychology , 61, 195–202. <http://dx.doi.org/10.1037/0022-3514.61.2.195>

Schwarz, N. and Clore, G.L. (1983) Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states . Journal of Personality and Social Psychology , 45, 513–523. <http://dx.doi.org/10.1037/0022-3514.45.3.513>

Schwarz, N. and Clore, G.L. (1996) Feelings and phenomenal experiences . In E.T. Higgins, ed. and A.W. Kruglanski (eds), Social Psychology. Handbook of Basic Principles , pp. 433–465. New York: Guilford Press.

Seibt, B., Häfner, M. and Deutsch, R. (2007) Prepared to eat: How immediate affective and motivational responses to food cues are influenced by food deprivation . European Journal of Social Psychology , 37, 359–379. <http://dx.doi.org/10.1002/ejsp.365>

Sloman, S.A. (1996) The empirical case for two systems of reasoning . Psychological Bulletin , 119, 3–22. <http://dx.doi.org/10.1037/0033-2909.119.1.3>

Solarz, A.K. (1960) Latency of instrumental responses as a function of compatibility with the meaning of eliciting verbal signs . Journal of Experimental Psychology , 59, 239–245. <http://dx.doi.org/10.1037/h0047274>

Sorabji, R. (2002) Emotion and Peace of Mind: From Stoic Agitation to Christian Temptation . Oxford: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780199256600.001.0001>

Strack, F. (1992) The different routes to social judgments: Experiential versus informational strategies . In L.L. Martin, ed. and A. Tesser (eds), The Construction of Social Judgment , pp. 249–275. Hillsdale, NJ: Erlbaum.

Strack, F. and Bless, H. (1994) Memory for nonoccurrences: Metacognitive and presuppositional strategies . Journal of Memory and Language , 33, 203–217. <http://dx.doi.org/10.1006/jmla.1994.1010>

Strack, F. and Deutsch, R. (2004) Reflective and impulsive determinants of social behavior . Personality and Social Psychology Review , 8, 220–247. http://dx.doi.org/10.1207/s15327957pspr0803_1

Strack, F. and Deutsch, R. (2007) The role of impulse in social behavior . In A.W. Kruglanski, ed. and E.T. Higgins (eds), Social Psychology. Handbook of Basic Principles , pp. 408–431. New York: Guilford.

Strack, F., Erber, R. and Wicklund, R.A. (1982) Effects of salience and time pressure on ratings of social causality . Journal of Experimental Social Psychology , 18, 581–594 <http://dx.doi.org/10.1016/0022-1031%2882%2990074-9>

- Strack, F. and Förster, J. (1998) Self-reflection and recognition: The role of metacognitive knowledge in the attribution of recollective experience . *Personality and Social Psychology Review* , 2, 111–123. http://dx.doi.org/10.1207/s15327957pspr0202_4
- Strack, F., Schwarz, N. and Gschneidinger, E. (1985) Happiness and reminiscing: The role of time perspective, affect, and mode of thinking . *Journal of Personality and Social Psychology* , 49, 1460–1469. <http://dx.doi.org/10.1037/0022-3514.49.6.1460>
- Strack, F., Werth, L. and Deutsch, R. (2006) Reflective and impulsive determinants of consumer behavior . *Journal of Consumer Psychology* , 16, 205–216. http://dx.doi.org/10.1207/s15327663jcp1603_2
- Topolinski, S. and Strack, F. (2009) The architecture of intuition: Fluency and affect determine intuitive judgments of semantic and visual coherence and judgments of grammaticality in artificial grammar learning . *Journal of Experimental Psychology. General* , 138, 39–63. <http://dx.doi.org/10.1037/a0014678>
- Tversky, A. and Kahneman, D. (1974) Judgment under uncertainty: Heuristics and biases . *Science* , 185, 1124–1131. <http://dx.doi.org/10.1126/science.185.4157.1124>
- Unkelbach, C., Forgas, J.P. and Denson, T.F. (2008) The turban effect: The influence of Muslim headgear and induced affect on aggressive responses in the shooter bias paradigm . *Journal of Experimental Social Psychology* , 44, 1409–1413. <http://dx.doi.org/10.1016/j.jesp.2008.04.003>
- Vohs, K.D. and Faber, R.J. (2007) Spent resources: Self-regulatory resource availability affects impulse buying . *Journal of Consumer Research* , 33, 537–547. <http://dx.doi.org/10.1086/510228>
- Whiteside, S.P. and Lynam, D.R. (2001) The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity . *Personality and Individual Differences* , 30, 669–689. <http://dx.doi.org/10.1016/S0191-8869%2800%2900064-7>

WiersR.W., RinckM., DictusM. and van den WildenbergE. (2009) Relatively strong automatic appetitive action-tendencies in male carriers of the oPRM1 G-allele . *Genes, Brain and Behavior* , 8, 101–106. <http://dx.doi.org/10.1111/j.1601-183X.2008.00454.x>

Wiers, R.W., Rinck, M., Kordts, R., Houben, K. and Strack, F. (2010) Retraining automatic action-tendencies to approach alcohol in hazardous drinkers . *Addiction* , 105, 279–287. <http://dx.doi.org/10.1111/j.1360-0443.2009.02775.x>

Wilson, T.D., Lindsey, S. and Schooler, T.Y. (2000) A model of dual attitudes . *Psychological Review* , 107, 101–126. <http://dx.doi.org/10.1037/0033-295X.107.1.101>

Winkielman, P. and Cacioppo, J.T. (2001) Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect . *Journal of Personality and Social Psychology* , 81, 989–1000. <http://dx.doi.org/10.1037/0022-3514.81.6.989>

Wittenbrink, B., ed. and Schwarz, N. (eds) (2007). *Implicit Measures of Attitudes: Procedures and Controversies* . New York: Guilford.

Wood, W. and Neal, D.T. (2007) A new look at habits and the habit-goal interface . *Psychological Review* , 843–863. <http://dx.doi.org/10.1037/0033-295X.114.4.843>

Zuckerman, M. (1994a) *Behavioral Expressions and Biosocial Bases of Sensation Seeking* . New York, NY: Cambridge University Press.

Zuckerman, M. (1994b) Impulsive unsocialized sensation seeking: The biological foundations of a basic dimension of personality . In J.E.W.T.D. Bates (ed.), *Temperament. Individual Differences at the Interface of Biology and Behavior* , pp. 219–255. Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10149-008>

Zuckerman, M. (1996) The psychobiological model for impulsive unsocialized sensation seeking: A comparative approach . *Neuropsychobiology* , 34, 125–129. <http://dx.doi.org/10.1159/000119303>

<http://dx.doi.org/10.4135/9781446249215.n6>

Handbook of Theories of Social Psychology: Volume 1

Construal Level Theory

Contributors: Yaacov Trope & Nira Liberman

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Construal Level Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n7>

Print pages: 118-135

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n7>

[p. 118 ↓]

Chapter 6: Construal Level Theory

Yaacov Trope, ed. and Nira Liberman, ed.

Abstract

We started with how the value of outcomes changes over temporal distance and ended up with what we hope is a step toward a general theory of psychological distance. There are different psychologies for the different dimensions of psychological distance – temporal, spatial, social, and hypotheticality. Without denying the uniqueness of each of these dimensions and more specific aspects comprising those dimensions, we believe they all entail transcending the present through constructing mental models of what is not directly experienced. The farther removed an object is from me on any distance dimension, the higher (more abstract) the level of mental construal of that object. Psychological distance thus expands or contracts depending on level of construal. Consistent with this proposal, research conducted in the framework of construal level theory (CLT) suggests that (1) different distance dimensions are mentally associated, (2) that distance on any of these dimensions influences and is influenced by higher levels of mental construal, and (3) that the various distances are, to some extent, interchangeable in their effects on prediction, evaluation, and choice.

Introduction

From a bird's eye view, social psychology stands out as a discipline that, more than any other discipline in the behavioral and social sciences, has sought to understand how people experience the immediate situation. The pursuit of this goal was inspired by Kurt

Lewin's field-theoretic emphasis on the critical role played by the forces operating in the "here-and-now." Following Lewin's lead, classic experiments in social psychology provide dramatic demonstrations of the power of the immediate situation and how individuals get caught up in it. For example, Asch's (1956) conformity experiment, Milgram's (1965) obedience studies, and Zimbardo's (1971) prison study vividly illustrate how individuals become engrossed in accommodating group pressure, formal authority, and social role requirements. People have always been known to conform, obey, and adhere to their social roles. The contribution of these and other classic social psychology experiments is in showing how readily individuals become mentally immersed in the intricacies of the immediate situation and how readily their pre-existing perceptions, attitudes, and values can be crowded out and no longer predict behavior. Consistent with this metatheoretical orientation, experimental social psychology has [p. 119 ↓] developed over the years an expertise, unique among the behavioral and social sciences, in how to mentally immerse people in the immediate situation.

Studying mental immersion in the present is important in its own right, but it is also important for another reason: It brings to the fore the question of what is the alternative psychological state. Bearing on this question, research and anecdotal evidence suggest that, perhaps unlike any other species, humans are sometime able to transcend the "here-and-now." They are able to recollect themselves in the past, plan the future, take others' perspective, cognize spatially remote places, and contemplate counterfactual alternatives to reality. In each case, a psychological distance from the self in the here-and-now is traversed. It is physically impossible to be in the past, future, or at different locations at the same time, as it is impossible to experience counterfactual alternatives to reality and other people's states. Yet humans have evolved an ability to broaden their spatial, temporal, and social horizons beyond the present to include the nonpresent. For example, participants in the Milgram obedience experiment might imagine how they would feel if they were the learner, how others would view what they are doing, what they would do if they were in a similar situation in the future, whether they have acted in the same way in the past or in other places, what would they do if the experimenter was not present, and more.

It is hard to tell how many participants actually had these thoughts in the Milgram experiment – probably a precious few; after all, the experiment was set up to make mental travel difficult. Were those who did engage in mental travel more likely than

others to disobey the experimenter's commands? Perhaps. We still don't know the answer, but this motivated our interest in the more general question of how mental travel occurs; that is, what are the psychological processes that enable people to mentally shift back and forth between being immersed in the here-and-now and removing oneself from it? This is the question that the construal level theory (CLT) of psychological distance is addressing.

A Team Narrative of the Theory Development

Temporal Construal

We started with a more specific, and seemingly simpler, question of how temporal distance of an outcome affects judgment and choice with respect to that outcome. This question was not new. For many years, researchers across different behavioral science disciplines, including psychology, economics, and political science, have studied how people make choices for their immediate future versus distant future (e.g., Ainslie, 1975; Elster, 1977; Kirby and Herrenstein, 1995; LaPiere, 1934; Loewenstein and Prelec, 1992; Lovallo and Kahneman, 2000; Mischel et al., 1969; Rachlin, 1995; Schelling, 1984) and how they evaluate their past (e.g., Gilovich and Medvec, 1995; Nisan, 1972; Ross, 1989). This research has yielded a wealth of findings on a wide range of phenomena, including time discounting, delay of gratification, shifts in level of aspiration, future planning, future optimism, overconfident prediction, regret, hindsight bias, and biased autobiographical memory.

We were particularly intrigued by two sets of phenomena: the “planning fallacy” and temporal discounting. The planning fallacy refers to the tendency, not unfamiliar to academics, to overcommit oneself when making plans for the future (Kahneman and Tversky, 1979). This phenomenon was brought to our attention by Mike Ross, who has been studying it at the time and who has been a great source of inspiration in our thinking about temporal construal. Mike Ross, Tom Gilovich, and their colleagues (e.g., Buehler et al., 1994; Gilovich et al., 1993) have reported ingenious experimental

demonstrations of people undertaking more activities than they [p. 120 ↓] can possibly accomplish and which they come to regret only when they get closer to the time of enactment. These findings strongly resonated with our own personal experience and that of colleagues marveling at themselves for repeatedly making the mistake of committing far in advance to writing book chapters, journal reviews, letters of recommendation, and the like. An exception, which might actually prove the rule, is Amos Tversky who, hearing about our interest, responded self-contentedly that he has firmly established with the administration at Stanford that he could not be asked to do anything with a deadline beyond six weeks.

The other set of phenomena, time discounting, is people's tendency to attach greater value to immediate outcomes than to delayed outcomes. This phenomenon has attracted our attention because it has been studied across many disciplines in the neural, behavioral, and social sciences. In fact, we could not (and still cannot) think of any single behavioral phenomenon that has been studied across such a broad range of disciplines as time discounting. We were particularly fascinated and inspired by Walter Mischel's many years of programmatic investigation of delay of gratification in children and the many insights into the nature of self-regulation that this research has inspired. Research on time discounting has focused on identifying factors that determine the discount rate. For example, affect-dependent time discounting suggests that affective outcomes undergo steeper time discounting than cognitive outcomes (Loewenstein, 1996; Loewenstein et al., 2001; Metcalfe and Mischel, 1999). Conflict theories (Lewin, 1951; Miller, 1944) assume that negative outcomes undergo steeper time discounting than do positive outcomes. Behavioral economists have shown that the discount rate depends on the magnitude of the value of outcomes, such that small rewards are discounted at a faster rate than large rewards (Green et al., 1997; Thaler, 1981).

Each of these lines of research employed well-developed methodologies and offered important descriptive and prescriptive implications for individuals and society. What seemed to be missing was an overarching framework that could integrate the diverse and sometime conflicting findings in the area as well as the specific hypotheses that have been offered to account for the findings. Our approach to this problem was based on the idea that people rely on schematic mental models for making judgments about their past and future (see Griffin and Ross, 1994; Kahneman and Lovallo, 1991; Ross, 1989). We thought that an integrative framework would have to take into account

how temporal distance from an object changes the way people mentally represent that object. The effects of temporal distance would then emerge from an active online construction of mental representations of future objects; hence the term “temporal construal.” Those effects, moreover, would involve a basic aspect of the way people represent objects but could nevertheless take place without conscious awareness, so that it would repeatedly lead people to change their mind as they get closer to the objects without insight into the source of their temporal inconsistency.

But what is that “basic aspect” of mental construal that changes with temporal proximity? Adopting a functional perspective, we asked what mental construal would be like in order for it to enable people to transcend the present and predict, evaluate, and take action with respect to the future. Our answer, which led to the development of CLT, built on past research in many areas of psychology, including memory,’ concept formation, categorization, causal reasoning, person perception, and goal hierarchies. We distinguished between high-level and low-level construals. High-level construals, we proposed, are schematic, decontextualized representations that extract the gist from the available information, emphasizing a few superordinate core features of events. Low-level construals, in contrast, are relatively unstructured, contextualized representations that include subordinate and incidental features of events. Consider, for example, two children playing [p. 121 ↓] catch with a ball in a backyard. A low-level construal of this activity might include such details as the age of the children, the color of the ball, and the temperature outside. In contrast, a high-level construal of this activity might simply be “having fun.” The high-level construal, “having fun,” disregards the unique features of the event, and involves an implicit decision about which features are central to the event and which are peripheral. Moving to high-level construal involves omitting features that are perceived as less central to the abstract construct in question.

In line with our functional approach, we viewed high-level construals as having evolved to represent distal events because, with distance, one needs to preserve the essential, invariant properties of the referent event. High-level construals are capable of fulfilling this function because they preserve the essential properties of a stimulus across momentary changes in appearance and through changes in time. High-level construals function somewhat like perceptual constancies, abstracting stimuli to their invariant properties. Low-level construals, by contrast, preserve a stimulus in minute detail, emphasizing its uniqueness rather than its similarity to other stimuli.

Based on this reasoning, and as described in more detail below, we launched a research program on temporal construal. Our first set of studies addressed the two phenomena mentioned earlier: the planning fallacy and time discounting. We started with a simple study asking students to indicate the number of weekly hours they planned to engage in each of several routine academic activities (e.g., "attend classes") and nonacademic activities ("sports and exercise") either next week or a week a few months later. Two results of this study caught our attention. One was that the total number of weekly hours planned for next week was considerably smaller than that planned for a week a few months later. The other was that the number of hours planned for different activities was negatively correlated for the near future week, but uncorrelated for the distant future week. Students seemed to make plans for the distant future as if there were no limits on their time resources and as if engaging in one activity didn't come at the expense of engaging in another.

We understood this and related planning fallacy phenomena as resulting from the tendency to construe distant actions in terms of their high-level, superordinate aspects rather low-level, subordinate features. Desirability concerns, which involve the action's end-state (i.e., the "why" aspect of the action) are superordinate aspects of actions, whereas feasibility concerns, which involve the means used to reach the end-state (i.e., the "how" aspect of the activity, the specific aspects of its enactment and its contextual constraints), are subordinate aspects of activities. CLT thus predicts that desirability concerns should receive greater weight over feasibility concerns as psychological distance increases. From a temporally distant perspective, activities are represented in terms of their desirability aspects, but as one gets closer in time to actual enactment, feasibility aspects become more prominent. The explanation of the overcommitment findings from our little weekly planning study seemed straightforward: temporal constraints and conflicts among different activities were feasibility aspects and therefore omitted from distant future plans. The time planned for each activity was based on its inherent attractiveness, irrespective of the time already allocated to other activities (Liberman and Trope, 1998).

The same logic applies to time discounting, the second issue we addressed early in our temporal construal research. We reasoned that different construals may entail different evaluations. For example, the construal "running subjects in the lab" may foster a less positive evaluation of an activity than the higher-level construal "conducting a

psychology study." Because CLT assumes that people use higher-level construals for distant future events than for near future events, it predicts that the value associated with low-level construals would be more prominent in evaluating near future events, whereas the [p. 122 ↓] value associated with high-level construals would be more prominent in evaluating distant future events. Time discounting, which has been commonly assumed to be a general principle in psychology, decision science, and economics, should hold only for the value associated with low-level construals. For value associated with high construals, the opposite may obtain, namely, time augmentation.

CLT predicts, then, that when the value associated with low-level construals is more positive than that associated with high-level construals, time delay would discount the attractiveness of an option. In contrast, when the value associated with low-level construals is less positive than that associated with high-level construals, time delay would augment the attractiveness of an option. Consistent with this prediction, we found that the value of the core, superordinate aspects of options were more influential in making a choice for the distant future, whereas the value of the secondary, subordinate aspects of options were more influential in making a choice for the near future (Trope and Liberman, 2000). For example, the extent to which an option promised to fulfill the individual's superordinate goal was more influential in choosing a distant future option, whereas incidental, goal-irrelevant considerations were more influential in choosing a near future option.

Psychological Distance

While conducting this research, we suspected that our conceptualization might be overly narrow. If, as we have assumed, high-level construals serve to transcend the present self and afford future time travel, couldn't they also serve to transcend the self in other respects? These construals may enable one to retrospect, imagine oneself in remote locations, take the perspective of other people, and contemplate counterfactual alternatives to reality. In each case, one traverses a psychological distance from oneself be it temporal distance (prospective and retrospective), spatial distance, social distance, or hypotheticality. High-level construal may expand not only one's temporal horizons, but also one's spatial horizons, social horizons, and the realm of possibility.

Correspondingly, low-level construals may enable one to become immersed in one's own direct experience of the here-and-now.

Switching between levels of construal may thus afford mental travel back and forth between the proximal and the distal across the various psychological distances. Time travel is extremely interesting and for many years has fascinated scholars across the social, cognitive, and neural sciences (e.g., Ainslie, 1975; Gilbert and Wilson, 2007; Schacter and Addis, 2007; Schelling, 1984). However, it might be only one instance of mental travel. As social psychologists, we were particularly intrigued by the idea that social distance might be thought of as one of the dimensions of psychological distance. Distinguishing between self and others, mine and yours, and taking others' perspective seemed fundamental to traversing psychological distance. The possession of this capability and its extension to dissimilar others, strangers, and outgroups, we thought may have played a key role in social life and in the development of civilization.

As social psychologists with interest in social cognition and self-regulation, we were also intrigued by the idea that the many distinctions that comprise what we call hypotheticality can be all thought of as aspects of psychological distance. These distinctions include real versus imaginary, factual versus counterfactual, true versus false, probable versus improbable, and expected versus unexpected. For example, perhaps people intuit improbable events as psychologically distant, as events one has to wait long or go far in order to encounter, and perhaps people employ high-level construals to incorporate such events into their judgments and choices.

In general, the functional approach we have taken suggests that the relationship between construal level and psychological [p. 123 ↓] distance is intrinsic. The ability to traverse psychological distance confers an obvious evolutionary advantage, and mental construal levels may have evolved, and may continue to evolve, in order to enable traversing increasingly greater psychological distances. This may be why level of construal and the expansion of mental horizons may have been linked throughout human evolution and why they appear to be linked in personal development. We felt that in relating CLT to these issues, we were responding to Lewin's (1951) call for developing a unified and general theory of psychological distance.

Questions Construal Level Theory Addresses

Thus far we have recounted how trying to understand specific time perspective effects has led us to develop a general CLT of psychological distance. But what are the new research questions the theory has led us to explore? What are its empirically testable hypotheses? There are three sets of questions that CLT has led us to ask: First, do all distances affect and are affected by level of construal? Second, are the various distances interrelated? That is, does distancing an object on one dimension lead people to expect the object to be distant on other dimensions? Third, do the various distances from an object similarly affect prediction, evaluation, and action regarding that object? The overarching, metatheoretical question was whether the level of generality of CLT is matched by the breadth of the phenomena this theory might predict. In this section we very briefly describe how we and our colleagues have addressed these questions.

Psychological Distance and Mental Construal

Our initial research examined the effect of temporal distance on the level of categorization of actions. However, to establish the generality of CLT, it was important to find out whether other aspects of construal, not only action categorization, were related to other psychological distances and whether the association was bidirectional. This led us to explore construal levels in perception, categorization, and inference. The following investigations illustrate these lines of research.

Perception

In a series of studies, participants completed what they believed to be sample items of a task that required abstraction of coherent images from fragmented or noisy visual input (the Gestalt Completion Test). Participants' performance improved when they

anticipated working on the actual task in the more distant future (Förster et al., 2004), when they thought the actual task was less likely to take place (Wakslak et al., 2006), and when social distance was enhanced by priming of high social status (Smith and Trope, 2006). A psychologically distant perspective thus seems to enable people to better see the “big picture.”

While abstraction improves the ability to perceive the gestalt in a visual array, it should have the opposite effect on performance when the task is to recognize missing elements within a gestalt. Distance should therefore have a detrimental effect on the ability to identify a missing element within a coherent whole. Indeed, participants did worse on sample items of this task when they believed they were less likely to later complete it (Wakslak et al., 2006).

Categorization

Our initial temporal construal research was based on the assumption that actions may be construed in high-level terms, which link them to a superordinate purpose (why one performs them), or in low-level terms, which link them to subordinate means (how one performs them). Consistent with this assumption, we found that participants tended to describe more distant future activities [p. 124 ↓] (e.g., studying) in high-level terms (e.g., “doing well in school”) rather than in low-level terms (e.g., “reading a textbook”) (Liberman and Trope, 1998). Similar effects emerged when actions were to take place in a spatially distant location (Henderson et al., 2006), when the actions were framed as unlikely to actually take place (Wakslak et al., 2006), and when the actor was described as dissimilar to the perceiver (Liviatan et al., 2008).

Going beyond action categorization, we then examined breadth of categorization of objects by asking participants to imagine an event (e.g., a camping trip) occurring in either the near or the distant future. For each event, participants grouped a set of related objects (e.g., tent, ball, snorkel) into as many groups as they deemed appropriate. Consistent with the idea that distance promotes the use of more abstract terms, participants who thought of a more distant event created fewer, broader groups of objects (Liberman et al., 2002). Reduced likelihood and social distance had the same effect (Smith and Trope, 2006; Wakslak et al., 2006). For example, objects that

pertained to less likely events (e.g., a trip that had a high probability of being canceled) were grouped into broader categories.

Bi-Directional Relationships between Construal Level and Distance

Do high-level construals, compared to low-level construals, lead individuals to think about psychologically more remote possibilities? Our research on this issue used a variety of manipulations of construal level (e.g., global versus local perceptual processing, generating categories versus exemplars, using broad versus narrow categories, identifying means versus ends, etc.). The results consistently showed that high-level construals lead individuals to think about more distant times, spatial locations, people, as well as relatively unlikely and rare events (Liberman and Förster, 2009; Wakslak and Trope, 2009).

The association between distance and level of construal was also demonstrated with implicit measures. In a series of studies using the Implicit Association Test (Bar-Anan et al., 2006 participants were presented with words from four categories: high-level construal (e.g., category names such as “drinks”), low-level construal (e.g., exemplar names such as “coke”), small psychological distance (e.g., socially proximal words such as “ours”, “friend”), and large psychological distance (e.g., socially distant words such as “theirs,” “stranger”). Participants mapped words from these four categories on two responses, pressing either a left key or a right key on the computer keyboard. On CLT-congruent trials, high-level stimuli were paired with distant stimuli and low-level stimuli were paired with proximal stimuli, whereas on CLT-incongruent trials high-level stimuli were paired with proximal stimuli and low-level stimuli were paired with distal stimuli. With all four dimensions of psychological distance, participants were faster with congruent than with incongruent pairings, suggesting that they implicitly associated psychological distance with high-level construal and psychological proximity with low-level construal.

It seems, then, that as psychological distance increases, construals become more abstract, and as level of abstraction increases, so too do perceptions of psychological

distance. This supports the basic tenet of CLT that abstract thinking is used to transcend the present and expand one's mental horizon by thinking farther into time and space and considering remote social targets and unlikely possibilities. At a metatheoretical level, we felt that the range of perceptual and cognitive phenomena CLT has led us to explore justified the level of generality of the theory.

The Association among Distance Dimensions

We have found that different distances are similarly related to construal and similarly. But the question still remained: do the distances in themselves have something in [p. 125 ↓] common? If the different distance dimensions have a shared meaning, as CLT contends, then these dimensions should be mentally associated. We felt that it is important for the theory to address this issue empirically. The idea was that different objects whose distance on one dimension is congruent with their distance on any other dimension would be associated even when they have little or nothing else in common. For example, remote locations would bring to mind the distant rather than the near future, other people rather than oneself, and unlikely rather than likely events.

Initial support for this hypothesis comes from a set of studies that assessed implicit association between spatial distance and other distance dimensions (Bar-Anan et al., 2007). Using a picture-word version of the Stroop paradigm (Stroop, 1935), participants discriminated between cues of one psychological distance dimension while ignoring cues of another psychological distance dimension. We predicted that it would be easier to perform the task when the relevant and irrelevant cues are congruent (as opposed to incongruent) in terms of psychological distance. Participants viewed perspective pictures containing an arrow that was pointing to either a proximal or distal point on the landscape shown in the picture; inside the arrow was printed a word denoting either a psychologically proximal entity (e.g., tomorrow, we, sure) or a psychologically distal entity (e.g., year, others, maybe). In a spatial discrimination version of the task, participants had to indicate whether the arrow pointed to a proximal or distal location. In a semantic discrimination version, participants had to indicate what the word printed on the arrow was. In both tasks, and across the distance dimensions, participants were

faster when responding to distant congruent stimuli (e.g., “we” printed on proximal arrow) than distant incongruent stimuli (e.g., “we” printed on distal arrow). The results suggest that people access the psychological distance of stimuli automatically, even when it is not directly related to people's current goals.

We concluded from these and related findings that the different psychological distance dimensions are associated and that psychological distance is an important aspect of meaning, common to spatial distance, temporal distance, social distance, and hypotheticality. It is also possible that cues of distance from events on one dimension affect the perceived distance from those events on other dimensions. For example, the spatial distance from an event may depend not only on its location relative to that of the perceiver but also on whether it is expected in the near or distant future, whether it occurred recently or a long time ago, whether it is probable or improbable, and whether it is expected to be experienced by oneself or another person. In this respect, the different psychological distances may be to some extent interchangeable.

The Socio-Psychological Applicability of Construal Level Theory

Having established that construals depend not only on the actual attributes of the objects but also on their psychological distance, we turned to the psychological consequences of this association. The question was whether all distances similarly affect (via their effect on construal) predictions, evaluation, and self-regulation. We have also started to explore the consequences of the distance-construal association for human relations.

Psychological Distance and Prediction

The function of high-level construals is to enable individuals to mentally transcend the here-and-now by forming a structured representation of the invariant features of the available information and projecting it onto distal objects. Consequently, predictions of future experiences would be more schematic [p. 126 ↓] than the actual experiences,

giving rise to a variety of prediction biases that stem from underweighting contextual and incidental features.

Two studies exemplify our research on this issue. One study, examining the effect of temporal distance on prediction of scientific findings, found that students were more confident that the exact same experiment would yield theory-confirming results when they expected the experiment to take place in a more distant point in time (Nussbaum et al., 2006). Apparently, when making predictions for the more distant experiments, participants gave more weight to the theory (high-level construal) and less weight to incidental noise factors (low-level construal). Another study investigated the effect of spatial distance on the tendency to base predictions on global rather than local information (Henderson and Fujita, 2006). New York University (NYU) participants viewed a series of graphs depicting information from the years 1999 to 2004 (e.g., average number of photocopies per student). The information was said to pertain to the NYU campus in Manhattan (spatially near condition) or to the NYU campus in Florence, Italy (spatially distant condition). Each graph showed either an upward or a downward trend, with the final year (2004) always deviating from that global trend. Participants estimated the likelihood that the year 2005 would be consistent with the general trend or with the more recent local deviation. In terms of CLT, global trends convey a high-level construal, whereas deviations, being local exceptions, should receive more weight in low-level construals. Consistent with this reasoning, spatial distance enhanced the tendency to predict on the basis of the global trend rather than the local deviation.

Psychological Distance and Evaluation

As noted earlier, a common assumption in the behavioral sciences is that the value of an outcome is discounted as temporal distance from the outcome increases. We proposed, in contrast, that temporal distance, as any psychological distance, should shift the overall attractiveness of an outcome closer to its high-level construal value and away from its low-level construal value. Two sets of studies addressed this issue.

The first set of studies examined desirability and feasibility concerns (Liberman and Trope, 1998). Desirability concerns involve the value of the action's end-state (a high-level construal), whereas feasibility concerns involve the means used to reach the end-

state (a low-level construal). Therefore, desirability concerns should receive greater weight over feasibility concerns as psychological distance increases. Consistent with this prediction, we found that as temporal distance from an activity (e.g., attending a guest lecture) increased, the attractiveness of the activity depended more on its desirability (e.g., how interesting the lecture was) and less on its feasibility (e.g., how convenient the timing of the lecture was). Similar results emerged with other psychological distance dimensions: probability (Todorov et al., 2007) and social distance (Liviatan et al., 2008). These findings suggest that distance increases the attractiveness of alternatives that are desirable but hard to obtain, but decreases the attractiveness of alternatives that are less desirable but easy to obtain. Extending this effect to the realm of risky choice, we found that people take higher risks (i.e., favoring bets with a low probability of winning a high amount over those that offer a high probability to win a small amount) in decisions about more distal bets (Sagristano et al., 2002).

The second set of studies concerned central and peripheral sources of value. According to CLT, central, goal-related features of outcomes constitute a high-level construal of these outcomes, whereas peripheral, goal-irrelevant features of outcomes constitute a low-level construal. Distancing an outcome should therefore increase the weight of central features relative to peripheral features. Support for this prediction was found in a [p. 127 ↓] study in which participants imagined buying a radio set in order to listen to morning programs either the next day or in one year (Trope and Liberman, 2000). In one version, participants read that the sound quality of the radio set was good, but that the clock that was incidentally included was relatively useless. In a different version, participants read that the sound quality of the radio set was poor, but that the clock turned out to be quite useful. As predicted, thinking about the radio set in the more distant future increased satisfaction when the sound quality was good and the clock was useless but decreased satisfaction when the sound quality was poor and the clock was good, indicating that time delay increased the weight of central features and decreased the weight of peripheral features. It seems, then, that people's overriding goals are more likely to guide their choices for psychologically distant than psychologically near situations.

Psychological Distance and Self-Regulation

Like predictions and evaluations, self-regulation should be increasingly based on high-level construal aspects as psychological distance increases. As outcomes seem more temporally, spatially, or socially remote or unlikely, actions should be guided more by one's central, global concerns and less by secondary, local concerns. Our research on self-control exemplifies our approach to this issue. Exercising self-control requires acting in line with ones' central, superordinate, and global considerations in the presence of more locally tempting alternatives. Because such considerations naturally relate to high-level construals, psychological distance should facilitate self-control. Indeed, people seem to be better able to choose delayed but valuable outcomes for the distant than for the near future (Loewenstein and Prelec, 1992). As another example, choosing a negative but diagnostic assessment of one's abilities rather than a flattering but nondiagnostic assessment requires prioritizing the long-term benefits of self-improvement over subordinate concerns about feeling good. Consistent with this prediction, participants were more likely to prefer the negative but diagnostic assessment when it was expected in the more distant future (Freitas et al., 2001). A recent series of studies has directly linked construal level to self-control, showing that forming a high-level construal of situations enables better self-control (e.g., choosing a delayed reward, enduring painful but valuable diagnostic procedures). In the same vein, research on delay of gratification in children showed that an abstract representation of the temptation increases delay relative to a more concrete representation (Fujita et al., 2006).

Human Relations across Psychological Distance

Much of our research has examined how people traverse psychological distance. Recently, we have started to explore what is arguably the most important aspect of this phenomenon, namely how people traverse psychological distances in their social relations. Many of the key issues in social psychology – how we connect with others,

how our sense of self is influenced by them, how we give to, take from, and reciprocate with other people – are inherently about crossing the gap between oneself in the here-and-now, and another person existing outside of that current direct experience. Our research on interpersonal similarity, social power, politeness, and social conflict illustrate the implications of CLT for social relations.

Interpersonal Similarity

In CLT, interpersonal similarity is a form of social distance (Liviatan et al., 2008). The less similar someone is to oneself, the more socially distant they typically seem; therefore, we hypothesized and found that behavior performed by a dissimilar other was represented at a higher level of construal than [p. 128 ↓] behavior performed by a similar other (Liviatan et al., 2008). Given this association between social similarity and mental construal, attraction to similar compared to dissimilar others should reflect the valence of low-level features more than the valence of high-level features. Indeed, Liviatan et al. (2008) showed that participants' liking of a similar (versus dissimilar) target was based to an increasing degree on the positivity of the target's behaviors and secondary help (i.e., low-level construals) over the positivity of the target's traits and primary help (i.e., high-level construals). It seems, then, that high-level construals expand our social horizons enabling us to relate to socially diverse people, whereas low-level construals guide our response to people who are like us.

Social Power

Social power may engender a sense of distance from others. Indeed, individuals who have power see themselves as less similar to and thus more distant from other people than individuals who have less power. This perception might be due to the fact that groups, organizations, and societies ordinarily have a pyramidal structure with fewer individuals occupying high-power positions than low-power positions. There is therefore greater similarity in the positions held by individuals with low power than individuals with high power. If social power makes people feel distant from others, then it should also predispose them to construe information abstractly, focus on the central aspects

of situations, disregard secondary aspects, and establish clear priorities. Consistent with this view, Smith and Trope (2006) research suggests that the distal perspective activated by the possession of social power promotes going beyond the information given, detecting the underlying structure, and abstracting from it superordinate, central features. Power-related construal may expand people's mental horizons, enabling them to transcend the immediate circumstances and take into account the past, future, a broad range of people, and unlikely possibilities. Do we think of individuals as suitable for power positions, or as actually holding such positions, when they articulate and enact high-level construals of the situation? Do we prefer power to be held by individuals who articulate plans that transcend the present and extend to the distant future, remote places, diverse groups, and unusual circumstances? These questions await future research.

Politeness

The theory of politeness (Brown and Levinson, 1987) views politeness as a signifier and producer of social distance. CLT views social distance as a specific case of psychological distance, and as such expects it to be related to level of construal and to other distances. Indeed, research by Stephan et al. (in press) shows that greater politeness is associated with higher-level construals and with greater temporal and spatial distance. For example, participants were more polite when they addressed a person whom they construed in terms of abstract goals and dispositions, when they expected the target to receive the message in the relatively distant future, and when they addressed individuals in relatively distant locations. Examining the opposite direction of influence, the research also showed that a request to generate polite statements prompted participants to use abstract verbs, that polite utterances were judged as pertaining to the relatively distant future and as directed to addressees in relatively remote locations.

These findings help to understand why different languages and cultures seem to signify politeness in similar ways. For example, standing very close to an interlocutor is generally considered to be less polite than standing farther away (i.e., greater spatial distance is associated with more politeness). Also, tense shifts in verbal communication tend to affect politeness, with the use of present tense being less polite

as compared to past or future tenses (i.e., greater temporal distance is associated with more politeness). Likewise, requests and remarks are considered more polite when presented in a more indirect, [p. 129 ↓] abstract form rather than a direct, concrete form (i.e., higher construal is associated with greater politeness). These examples demonstrate that greater politeness may be achieved in social relations and communication when they entail greater spatial or temporal distancing and abstraction. This is neither a coincidence nor an arbitrary convention, but rather an expression of an underlying regularity wherein politeness, as a regulator of social distance, is expressed by a high level of construal and greater distance on various dimensions. Of course, different cultures might use different distance and construal cues to convey politeness (e.g., the use of plural as a polite way of addressing a person does not exist in many languages). However, CLT predicts that such cues would always use higher-level construals and greater distances to communicate greater politeness, rather than less politeness. This prediction awaits empirical cross-cultural examination.

Social Conflict

Issues within an interpersonal negotiation can differ in their centrality and worth. If a pair of negotiators can trade off their lowest and highest priority issues (e.g., give in on secondary issues in exchange for getting what they want on high-priority issues; a process called logrolling), they are more likely to succeed in “expanding the pie,” maximizing both individual and joint outcomes. Because negotiators should be expected to focus more on central concerns and less on peripheral concerns as distance increases, we would expect to see more logrolling agreements in a distant future than near future context. Examining this idea within the context of a live negotiation, Henderson et al. (2006) found that while 91 percent of dyads with a temporally distant perspective reached a fully logrolling agreement, only 50 percent of dyads with a temporally near perspective did so. The enhanced reciprocal concessions made by dyads with the temporally distant perspective culminated in better negotiated individual and joint outcomes. Moreover, research on the role of construal levels in the negotiation process has shown that negotiators who construed issues abstractly rather than concretely were more likely to discover integrative agreements (Henderson and Trope, 2009).

These findings have implications for how negotiators handle conflicts over minor versus major issues and specific interests versus broad values and ideological differences in situations that do not allow for trade-offs (Harinck and De Dreu, 2004). Specifically, CLT suggests that the resolution of minor issues and specific interests should be hindered when individuals adopt a psychological proximal perspective or a lower-level construal and facilitated by having a more distal perspective and abstract construal. Conversely, the resolution of conflict over major issues, values, and ideological differences should be hindered when individuals adopt psychological distant perspective or a higher-level construal and facilitated by a less abstract construal and a psychologically proximal perspective.

In summary, a range of studies suggests that people rely on high-level construals to a greater extent when predicting, evaluating, and taking action with respect to more distant situations. Once again, the broad range of distance-dependent responses uncovered by this line of CLT research seemed to justify the level of generality at which the theory was formulated. Moreover, because these responses involved potentially consequential judgments and choices in the social context, not only construals, the scope of the findings we have obtained is testimony to the practicality of the theory and its applicability to the functioning of individuals in groups and society.

What Construal Level Theory is Not

In the preceding two sections, we have discussed the questions addressed by CLT. In this final section we discuss questions that [p. 130 ↓] one might address to CLT. Throughout the history of social psychology, three questions have often been asked of theories of social cognitive processes: Do the processes lead to accurate judgments about reality? Do they facilitate goal attainment? Do they promote social accommodation? Applied to CLT's distinction between high-level and low-level mental construals, these questions become: Does one construal level lead to more accurate judgments than the other? Is one construal level more conducive to goal attainment? Is one construal level more conducive to social accommodation? We consider these three questions in turn.

Is One Construal Level More Conducive to Accurate Predictions than the Other?

It is debatable whether there is a single objective criterion for accuracy. But supposing there is some verifiable event – in one's personal life or in the social or physical reality – the question is which construal process leads to more accurate prediction of that event. Our answer is: it depends. In reality, an event might be determined by low-level or high-level factors. Quite obviously, then, when high-level factors determine an event, high-level construals are likely to be more accurate, and when low-level factors determine an event, low-level construals are likely to be more accurate. In other words, any given level of construal would lead to more accurate predictions of future events to the extent that it matches the level of factors that actually determine the occurrence of those events.

Of particular interest are predictions of one's own evaluative, emotional, or behavioral reaction to events (e.g., evaluations of a course, job, or vacation). Research on affective forecasting, for example, has shown that people's predictions of their affective reactions to a future event are inaccurate (Wilson and Gilbert, 2003). We think that this is, in part, because in affective forecasting studies the levels of construal underlying people's predictions and their actual reactions do not match. Specifically, reactions are assessed soon after the event occurs and are therefore based on a low-level construal of the event. In contrast, predictions of one's reactions to an event are made from a more distant perspective and therefore likely to be based on high-level construals of the event. The mismatch between the construal levels underlying reactions and predictions should make the predictions inaccurate, as affective forecasting research shows. For example, a scenic (high-level aspect) but uncomfortable (low-level aspect) bike trip is likely to seem more enjoyable in prospect than during the trip.

However, although the construal levels of predictions and *immediate* reactions to an event are mismatched, the construal levels of predictions and subsequent *retrospective* reactions to the event might be matched, rendering predictions relatively more accurate. For example, both predictions and distant retrospective reactions are likely to be based on ends rather than means to those ends. To use the bike trip example, both

prospective and retrospective evaluations of a scenic but uncomfortable bike ride are likely to be positive, and predictions of retrospective evaluations would thus be relatively accurate. In short, failure to predict one's immediate reactions to a future event is one of many possibilities: predictions from a distal perspective may better predict long-term than short-term reactions to events, whereas predictions from a proximal perspective may better predict short-term than long-term reactions to events.

Is One Construal Level More Conducive to Goal Attainment than the Other?

Again, our answer to this question is that it depends. First, in reality, the attainment of a goal may depend on high-level or low-level factors. Therefore, high-level or low-level construals are more conducive to the attainment [p. 131 ↓] of a goal if they match the level of factors that actually determine the attainment of that goal. For example, abstract thinking would promote the attainment of goals that require detection of global patterns, whereas concrete thinking would promote the attainment of goals that require detection of specific details (see Wakslak et al., 2006). Second, and perhaps more interesting, the goal itself may be a high-level or low-level goal. High-level goals are superordinate, global, and central, whereas low-level goals are subordinate, local, and secondary. Therefore, high-level construals and a distal perspective might promote high-level goal pursuit (and thus also self-control), whereas low-level construals and a proximal perspective might promote low-level goal pursuit. For example, abstract thinking would promote pursuit of one's high-priority goals, whereas concrete thinking would promote pursuit of one's low-priority goals (see Henderson et al., 2006; Eyal et al., 2009).

Is One Construal Level More Conducive to Social Accommodation than the Other?

Yet again, the answer is that it depends. Taking the perspective of others is a form of traversing psychological distance and therefore likely to be facilitated by high-level construals. Taking others' perspective, in turn, promotes forming a socially shared

reality and might therefore also promote social alignment with the group opinion. However, the group opinion might be entirely incidental to the decision the individual is facing. For example, the political opinions voiced by strangers one happens to meet in a local bar are likely incidental to one's decision how to vote in an upcoming election. Individuals at a high level of construal are therefore less likely to align their opinion to that of this group than are individuals at a low level of construal. For example, when the election is temporally or spatially distant, thus activating a high-level construal of the vote, individuals are likely to resist the group opinion and, instead, rely on their own preexisting attitudes. However, when the election is temporally or spatially proximal, thus activating low-level construals, individuals' opinions are more likely to shift toward that of the group. In general, a person's judgments are more likely to conform to those espoused by an incidental group when psychological distance is small rather than large. Recent research on attitude alignment with incidental strangers provides support for this hypothesis (Ledgerwood et al., 2009). For example, participants' voting intentions aligned with those of an incidental stranger when the policy was going to be implemented in the near future. In contrast, participants were unaffected by the incidental stranger's views when the policy was going to be implemented in the distant future. Moreover, this pattern was replicated when construal (abstract versus concrete) was directly manipulated via a mindset induction.

Conclusion

We started with how the value of outcomes changes over temporal distance and ended up with what we hope is a step toward a general theory of psychological distance. There are different psychologies for the different dimensions of psychological distance – temporal, spatial, social, and hypotheticality. Examples are the psychologies of time perspective, propinquity, social perspective taking, ingroup and outgroup perception, counterfactual thinking, and source monitoring. Without denying the uniqueness of each of these dimensions and the even more specific aspects comprising those dimensions, we believe they all entail transcending the present through constructing mental models of what is not directly experienced. The farther removed an object is from me on any distance dimension, the higher (more abstract) the level of mental construal of that object. Psychological distance thus expands or [p. 132 ↓] contracts depending on level

of construal. Consistent with this proposal, research conducted in the framework of CLT suggests that (1) different distance dimensions are mentally associated, (2) distance on any of these dimensions influences and is influenced by higher levels of mental construal, and (3) the various distances are, to some extent, interchangeable in their effects on prediction, evaluation, and choice.

Let us note in conclusion that CLT is a theory of cognition, motivation, and self-regulation. Levels of construal refer not only to how people represent and process information about objects, it also refers to the goals people want to attain and the plans they make to actually attain them. High-level construals entail pursuing superordinate goals (e.g., being healthy), whereas low-level construals involve pursuing subordinate goals (e.g., avoiding fatty food). High-level construals entail prioritizing one's primary and global goals over secondary and local ones, whereas low-level construals do not sharply distinguish between these types of goals. High-level construals entail forming general and structured plans for achieving one's goals, whereas low-level construals entails forming contextualized and ad-hoc plans. Most important, high-level construals enable us to care about and act upon what is not present, whereas low-level construals enable us to care about and act upon what is present.

References

- Ainslie, G. (1975) Specious reward: A behavioral theory of impulsiveness and impulse control . Psychological Bulletin , 82, 463–496. <http://dx.doi.org/10.1037/h0076860>
- Asch, S.E. (1956) Studies of independence and conformity: A minority of one against a unanimous majority . Psychological Monographs , 70, Whole No. 416.
- Bar-Anan, Y., Liberman, N., Trope, Y. and Algom, D. (2007) Automatic processing of psychological distance: Evidence from a Stroop task . Journal of Experimental Psychology. General , 136, 610–622. <http://dx.doi.org/10.1037/0096-3445.136.4.610>
- Bar-Anan, Y., Liberman, N. and Trope, Y. (2006) The association between psychological distance and construal level: Evidence from an Implicit Association

Test , Journal of Experimental Psychology. General , 135, 609–622. <http://dx.doi.org/10.1037/0096-3445.135.4.609>

Brown, P. and Levinson, S. (1987) Politeness: Some universals in language usage . Cambridge, England: Cambridge University Press.

Buehler, R., Griffin, D. and Ross, M. (1994) Exploring the ‘planning fallacy’: Why people underestimate their task completion times . Journal of Personality and Social Psychology , 67, 366–381. <http://dx.doi.org/10.1037/0022-3514.67.3.366>

Elster, J. (1977) Ulysses and the sirens: A theory of imperfect rationality . Social Science Information/sur les sciences sociales , 16, 469–526. <http://dx.doi.org/10.1177/053901847701600501>

Eyal, T., Liberman, N., Sagristano, M. and Trope, Y. (2009) Resolving value conflicts in making choices for the future . Unpublished manuscript, Ben Gurion University, Israel.

Förster, J., Friedman, R.S. and Liberman, N. (2004) Temporal construal effects on abstract and concrete thinking: Consequences for insight and creative cognition . Journal of Personality and Social Psychology , 87, 177–189.

Freitas, A.L., Salovey, P. and Liberman, N. (2001) Abstract and concrete self-evaluative goals . Journal of Personality and Social Psychology , 80, 410–412. <http://dx.doi.org/10.1037/0022-3514.80.3.410>

Fujita, K., Trope, Y., Liberman, N. and Levin-Sagi, M. (2006) Construal levels and self-control . Journal of Personality and Social Psychology , 90, 351–367. <http://dx.doi.org/10.1037/0022-3514.90.3.351>

Gilbert, D.T. and Wilson, T.D. (2007) Prospection: Experiencing the future . Science , 317, 1351.

Gilovich, T. and Medvec, V.H. (1995) The experience of regret: What, when, and why . Psychological Review , 102, 379–395. <http://dx.doi.org/10.1037/0033-295X.102.2.379>

Gilovich, T., Kerr, M. and Medvec, V.H. (1993) Effect of temporal perspective on subjective confidence . *Journal of Personality and Social Psychology* , 64, 552–560. <http://dx.doi.org/10.1037/0022-3514.64.4.552>

Green, L., Meyerson, J. and McFadden, E. (1997) Rate of temporal discounting decreases with amount of reward . *Memory and Cognition* , 25, 715–723. <http://dx.doi.org/10.3758/BF03211314>

Griffin, D.W. and Ross, L. (1994) Subjective construal, social inference, and human misunderstanding . In M. Zanna (ed.) *Advances in Experimental Social Psychology* , 26, 319–360. New York: Academic Press.

Harinck, F. and De Dreu, C.K.W. (2004) Negotiating interests or values and reaching integrative agreements: The importance of time pressure and temporary impasses . *European Journal of Social Psychology* , 34, 595–612. <http://dx.doi.org/10.1002/ejsp.218>

Henderson, M.D., Fujita, K., Trope, Y. and Liberman, N. (2006) Transcending the ‘here’: The effect of spatial distance on social judgment . *Journal of Personality and Social Psychology* , 91, 845–856. <http://dx.doi.org/10.1037/0022-3514.91.5.845>

Henderson, M.D., Trope, Y. and Carnevale, P. (2006) Negotiation from a near and distant time perspective . *Journal of Personality and Social Psychology* , 91, 712–729. <http://dx.doi.org/10.1037/0022-3514.91.4.712>

Henderson, M.D. and Trope, Y. (2009) The effects of abstraction on integrative agreements: When seeing the forest helps avoid getting tangled in the trees . *Social Cognition* , 27, 399–415. <http://dx.doi.org/10.1521/soco.2009.27.3.402>

Kahneman, D. and Lovallo, D. (1991) Timid choices and bold forecasts: A cognitive perspective on risk taking . *Management Science* , 39, 17–31. <http://dx.doi.org/10.1287/mnsc.39.1.17>

Kahneman, D. and Tversky, A. (1979) Prospect theory: An analysis of decisions under risk . *Econometrica* , 47, 263–291. <http://dx.doi.org/10.2307/1914185>

Kirby, K.N. and Herrnstein, R.J. (1995) Preference reversals due to myopic discounting of delayed reward . *Psychological Science* , 6, 83–89. <http://dx.doi.org/10.1111/j.1467-9280.1995.tb00311.x>

LaPiere, R.T. (1934) Attitudes versus actions . *Social Forces* , 13, 230–237. <http://dx.doi.org/10.2307/2570339>

Ledgerwood, A., Trope, Y. and Chaiken, C. (2009) Attitudes in their social context: Malleability, stability, and the role of construal . Unpublished manuscript, University of California, Davis.

Lewin, K. (1951) Field Theory in Social Science . New York: Harpers

Liberman, N. and Förster, J. (2009) Distancing from experienced self: How global versus local perception affects estimation of psychological distance . *Journal of Personality and Social Psychology* , 97, 203–216. <http://dx.doi.org/10.1037/a0015671>

Liberman, N. and Trope, Y. (1998) The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory . *Journal of Personality and Social Psychology* , 75, 5–18. <http://dx.doi.org/10.1037/0022-3514.75.1.5>

Liberman, N., Sagristano, M. and Trope, Y. (2002) The effect of temporal distance on level of construal . *Journal of Experimental Social Psychology* , 38, 523–535. <http://dx.doi.org/10.1016/S0022-1031%2802%2900535-8>

Liviatan, I., Trope, Y. and Liberman, N. (2008) Interpersonal similarity as a social distance dimension: Implications for perception of others' actions . *Journal of Experimental Social Psychology* , 44, 1256–1269. <http://dx.doi.org/10.1016/j.jesp.2008.04.007>

Liviatan, I., Trope, Y. and Liberman, N. (2010) A construal level analysis of the relationship between similarity and attraction . Unpublished manuscript. New York University.

- Loewenstein, G.F. (1996) Out of control: Visceral influences on behavior . *Organizational Behavior and Human Decision Processes* , 65, 275–292. <http://dx.doi.org/10.1006/obhd.1996.0028>
- Loewenstein, G. and Prelec, D. (1992) Anomalies in intertemporal choice: Evidence and an interpretation . *Quarterly Journal of Economics* , 107, 53–587. <http://dx.doi.org/10.2307/2118482>
- Loewenstein, G.F., Weber, E.U., Hsee, C.K. and Welch, N. (2001) Risk as feelings . *Psychological Bulletin* , 127, 267–286. <http://dx.doi.org/10.1037/0033-2909.127.2.267>
- Lovallo, D. and Kahneman, D. (2000) Living with uncertainty: Attractiveness and resolution timing . *Journal of Behavioral Decision Making* , 13, 179–190. <http://dx.doi.org/10.1002/%28SICI%291099-0771%28200004/06%2913:2%3C179::AID-BDM332%3E3.0.CO;2-J>
- Metcalfe, J. and Mischel, W. (1999) A hot/cool-system analysis of delay of gratification: Dynamics of willpower . *Psychological Review* , 106, 3–19. <http://dx.doi.org/10.1037/0033-295X.106.1.3>
- Milgram, S. (1965) Some conditions of obedience and disobedience to authority . *Human Relations* , 18, 57–76. <http://dx.doi.org/10.1177/001872676501800105>
- Miller, N.R. (1944) Experimental studies of conflict . In McV. Hunt (ed.) *Personality and the Behavior Disorders* , pp. 431–465. New York: Ronald Press.
- Mischel, W., Grusec, J. and Maters, J.C. (1969) Effects of expected delay time on the subjective value of rewards and punishments . *Journal of Personality and Social Psychology* , 11, 363–373. <http://dx.doi.org/10.1037/h0027265>
- Nisan, M. (1972) Dimension of time in relation to choice behavior and achievement orientation . *Journal of Personality and Social Psychology* , 21, 175–182. <http://dx.doi.org/10.1037/h0032245>

Nussbaum, S., Liberman, N. and Trope, Y. (2006) Predicting the near and distant future . *Journal of Experimental Psychology: General* , 135, 152–161. <http://dx.doi.org/10.1037/0096-3445.135.2.152>

Rachlin, H. (1995) Self control: Beyond commitment . *Behavioral and Brain Sciences* , 18, 109–159. <http://dx.doi.org/10.1017/S0140525X00037602>

Ross, M. (1989) Relation of implicit theories to the construction of personal histories . *Psychological Review* , 96, 341–357. <http://dx.doi.org/10.1037/0033-295X.96.2.341>

Sagristiano, M.D., Trope, Y. and Liberman, N. (2002) Time-dependent gambling: Odds now, money later . *Journal of Experimental Psychology: General* , 131, 364–536. <http://dx.doi.org/10.1037/0096-3445.131.3.364>

Schacter, D.L. and Addis, D.R. (2007) The cognitive neuroscience of constructive memory: Remembering the past and imagining the future . *Philosophical Transactions of the Royal Society* , 362, 773–786. <http://dx.doi.org/10.1098/rstb.2007.2087>

Schelling, T. (1984) Self command in practice, in theory and in a theory of rational choice . *American Economic Review* , 74, 1–11.

Smith, P.K. and Trope, Y. (2006) You focus on the forest when you're in charge of the trees: Power priming and abstract information processing . *Journal of Personality and Social Psychology* , 90, 578–596 <http://dx.doi.org/10.1037/0022-3514.90.4.578>

Stephan, E., Liberman, N. and Trope, Y. (2010) Politeness and social distance: A construal level perspective . *Journal of Personality and Social Psychology* , 98, 268–80. <http://dx.doi.org/10.1037/a0016960>

Stroop, J.R. (1935) Studies of interference in serial verbal reactions . *Journal of Experimental Psychology: General* , 18, 643–662. <http://dx.doi.org/10.1037/h0054651>

Thaler, R.H. (1981) Some empirical evidence on dynamic inconsistency . *Economic Letters* , 8, 201–207. <http://dx.doi.org/10.1016/0165-1765%2881%2990067-7>

Todorov, A., Goren, A. and Trope, Y. (2007) Probability as a psychological distance: Construal and preferences , *Journal of Experimental Social Psychology* , 43, 473–482. <http://dx.doi.org/10.1016/j.jesp.2006.04.002>

Trope, Y. and Liberman, N. (2000) Temporal construal and time-dependent changes in preference . *Journal of Personality and Social Psychology* , 79, 876–889. <http://dx.doi.org/10.1037/0022-3514.79.6.876>

Wakslak, C.J., Trope, Y., Liberman, N. and Alony, R. (2006) Seeing the forest when entry is unlikely: Probability and the mental representation of events . *Journal of Experimental Psychology: General* , 135, 641–653. <http://dx.doi.org/10.1037/0096-3445.135.4.641>

Wakslak, C.J. and Trope, Y. (2009) The effect of construal-level on subjective probability estimates . *Psychological Science* , 20, 52–58. <http://dx.doi.org/10.1111/j.1467-9280.2008.02250.x>

Wilson, T.D. and Gilbert, D.T. (2003) Affective forecasting . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 35, 345–411. San Diego: Academic Press.

Zimbardo, P. G. (1971) The power and pathology of imprisonment . *Congressional Record* . (Serial No. 15, 1971–10–25). Hearings before Subcommittee No. 3, of the Committee on the Judiciary, House of Representatives, Ninety-Second Congress, First Session on Corrections, Part II, Prisons, Prison Reform and Prisoner's Rights: California . Washington, DC: U.S. Government Printing Office.

<http://dx.doi.org/10.4135/9781446249215.n7>

Handbook of Theories of Social Psychology: Volume 1

An Attribution Theory of Motivation

Contributors: Bernard Weiner

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "An Attribution Theory of Motivation"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n8>

Print pages: 135-156

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n8>

[p. 135 ↓]

Chapter 7: An Attribution Theory of Motivation

Bernard Weiner, ed.

Abstract

Expectancy value theory as modified by Atkinson, locus of control research initiated by Rotter, and the naïve psychology of achievement performance advanced by Heider provided the stepping stones for the formulation of the attribution theories of motivation presented in this chapter. Two theories are reviewed – an intrapersonal conception that primarily addresses achievement striving, and an interpersonal theory that focuses on social behaviors including help-giving, aggression, and reactions to the stigmatized. Causal beliefs, their underlying properties, and distinctive links to emotions form the foundation for these theories, which suggest a thinking-feeling-action sequence as the “deep structure” for motivated behavior. Application of the theories to achievement striving, physical and mental health, and social support are discussed.

Introduction

To understand the particular version of attribution theory presented in this chapter, I must first introduce my mentor, John W. Atkinson. Atkinson took part in the pursuit of a “grand theory” of motivation. This goal characterized the study of motivation for about half a century, from roughly 1930 to 1975. For the experimental motivation psychologist, the aim was to identify the determinants of action and specify their mathematical

and/or sequential (temporal) relations. The most influential of these approaches was associated with Clark Hull and Kenneth Spence (Hull, 1943; Spence, 1956), the prime creators of drive theory. Their conception specified that behavior is determined by Drive \times Habit and other factors such as incentives. A competing theoretical approach was linked to Edward Tolman (1932), Julian Rotter (1954), and Atkinson (1957), who instead argued that behavior is directed by Expectancy \times Value and, for Atkinson, motives as well. Kurt Lewin (1938), who also may be considered an Expectancy \times Value theorist, additionally proposed a temporal link connecting the behavioral determinants. He postulated [p. 136 ↓] that an object acquires motivational properties (a positive or negative incentive valence) only after there is a need within the organism. This results in a motivation sequence of: need \rightarrow incentive \rightarrow force (behavior tendency).

These grand theories were presumed to account for all behavior, regardless of content domain. If not content-free, then at least the conceptions could be called content independent (although their experimental instantiation was often in quite narrow settings). As an example, for Expectancy/Value theorists motivated behavior across contexts is a product of what one expects to receive multiplied by the perceived probability of attaining that goal, with motivated behavior conceived as a rational, hedonic choice. Hence, the theories are called “grand,” perhaps a shortened version of grandiose, for their aspirations were formidable and their hopes embraced a unifying theory for all behavior. It is reasonable to propose that Einstein was the guiding light, with recognition of some limitations imposed by animate as opposed to inanimate objects of study. Indeed, while I was a graduate student, Atkinson was immersed in Isaac Newton and attempted to incorporate Newton's principles of motion into motivation theory (see Atkinson and Birch, 1970).

My use of attribution-related concepts, which I impose on issues in motivation rather than (or, in addition to) social perception, imitates the grand theory pursuits. I attempt to specify the determinants of behavior (as the Drive and Expectancy/Value theorists) and identify their sequential arrangement (as did Lewin), albeit with causal beliefs as the theoretical bedrock. This theoretical approach distinguishes my version of attribution theory from other attribution theories and from many current motivation and social psychological theories as well. That is primarily because the theoretical structure includes multiple components in specified interrelations and transcends specific content domains, which I regard as the *sine qua non* for the label of “theory.”

There is another historical antecedent, most clearly associated with E.L. Thorndike and his law of effect, which guides my work. Thorndike (1911) stated that behaviors previously rewarded will be repeated, whereas those that were punished will be avoided or extinguished. In this manner, Thorndike incorporated the past into the present. As an attribution theorist, I also ask how the past makes its way into the present. But rather than having reinforcement as the mechanism, it is contended that the interpretation of the past, that is, the perceived causes of prior events, determine what will be done in the future. Why one was rewarded or punished, not the outcome per se, directs the future, so that both reward and punishment may have positive or negative motivational consequences.

Theoretical Beginnings

Atkinson's version of Expectancy/Value theory embraced three idiosyncratic components:

- 1. As already revealed, motivation is also determined by motives, so that the equation for strength of motivation is Motive \times Expectancy \times Value. The motive Atkinson was wed to was the need for achievement.
- 2. In achievement settings only, incentive (value) was conceptualized as an affect, pride in accomplishment (here I consider only approach motivation and positive affect).
- 3. In achievement settings only, incentive was postulated to be inversely related to expectancy so that pride was presumed to be greater given success at a difficult task (low expectancy of success) rather than at an easy task (high expectancy of success).

A number of perils awaited as I continued in this research tradition following graduate school. First, achievement needs were assessed with a projective instrument, the Thematic Apperception Test (TAT). Scoring and motive classification decisions were time-consuming. Then, those designated as [p. 137 ↓] high or low in achievement needs were called back for an experimental manipulation, typically involving success or failure at tasks varying in subjective expectancy of success. But by the time I completed the individual difference assessment, some participants had fulfilled their experimental

requirements. Others would not take part in the second phase of the research. One consequence of this procedure and the loss of subjects was that I could only complete a few experiments yearly, hardly a fountain of empirical research for an assistant professor.

Furthermore, the theory was very restricted in its predictions, in spite of the “grand” foundation. The main theoretically generated hypothesis was that individuals high in achievement needs are especially attracted to intermediate difficulty tasks relative to persons low in achievement needs. Unfortunately, this and other hypotheses related to motive-group differences often were not confirmed in experimental studies (see Weiner, 1992). In short, I had reason to fear the outcome of any later tenure decision.

I therefore began to search for other predictors and predictions and more economical experimental procedures that would facilitate my research output. About that time, Julian Rotter (1966) published a monograph examining “locus of control.” I was previously unaware of this body of work, which grew from studies manipulating skill versus luck tasks and their influence on expectancies. It seemed logical to pursue the idea that persons with high achievement needs view the world as controllable by them, more so than persons low in achievement needs. That is, persons high in achievement needs attribute outcomes to internal factors such as skill, whereas those low in achievement needs perceive luck (external control) to be the major determinant of their success and failure. I devised a correlation study to test these ideas, administering achievement and locus of control measures to a large subject population (Weiner and Potepan, 1970). This study did yield some positive findings. I find it amusing, or frightening, that two measures (the TAT and locus of control scale) and two constructs (need for achievement and locus of control) that I now reject launched my research direction in attribution theory. I am unsure if scientific progress is better made by standing on the shoulders of giants to see further or by jumping off these giants to attain a different perspective.

After examining the locus of control construct and measure, a few enigmas became evident. For example, if one fails because of a perceived lack of aptitude, the locus is internal yet the cause is not subject to volitional change. That is, there can be internality without control. In addition, if one succeeds at a task because of perceived high ability, then subsequent failure at that task would not be ascribed to low ability. But this is

not the case given effort as the perceived cause – success or failure at the same task could be regarded as due to the degree of effort expenditure. Furthermore, the perceived determinants of success are not the same as failure – evidence of personal enhancement is pervasive, so that success more than failure elicits beliefs that the self played a role. In sum, distinctions between locus and control, ability and effort, and success and failure have to be made when considering causal ascriptions. These conclusions promoted my skepticism that locus of control is a trait and/or a unitary construct (just as I questioned whether the achievement motive is a trait inasmuch as one could be motivated, e.g., at tennis but not ping-pong). It also pushed me to think harder about perceived causality, the gateway to attribution theory. And it raised a lifelong question that is implicit in all my work, namely: “In what ways do ability and effort differ?”

I was mulling over these issues when reading a book by Richard de Charms (1968) entitled *Personal Causation*. De Charms offered a distinction quite similar to Rotter, labeling some individuals “origins” and others “pawns.” In his book he also reviewed an experiment by Schmitt (1964) in which [p. 138 ↓] moral judgments are made about individuals who do not repay their debts either because of lack of ability to work (e.g., because of illness) or insufficient effort (e.g., not working because of laziness). I recognized that ability and effort also could be varied as causes of success and failure in an achievement context. In a simulation experiment (Weiner and Kukla, 1970), we described schoolchildren as succeeding or failing and factorially varied whether they had ability or not, and exerted effort or not. The participants were instructed to evaluate (reward or punish) these students. These studies were easy to conduct, the variables easy to manipulate, the results easy to evaluate, and the findings systematic and replicable with no individual differences involved. Among the results, the data revealed that the pattern of low ability, high effort, and success produces the greatest reward, whereas the combination of high ability, low effort, and failure gives rise to the greatest punishment. Hence, lack of ability is a positive facilitator of achievement evaluation (holding outcome and effort constant), and achievement evaluation shares characteristics with other moral judgments. I felt more at ease regarding my future empirical prospects.

I submitted a manuscript of about 20 pages for publication, containing three experiments, to the *Journal of Personality and Social Psychology*. The editor at that

time was extremely critical, insightful, smart, and wordy. He wrote a ten-page single-spaced editorial response, with an invitation to resubmit. By the time I read and understood his comments, I had conducted a fourth experiment, which was included in the resubmission. This was answered with about a five-page editorial response, again asking for a resubmission; I resent a manuscript with five experiments. The editor responded in his usual manner, this time with an abbreviated two-page comment. Finally, a six-experiment article was accepted and published (Weiner and Kukla, 1970). It was a key turning point in my research directions.

Theoretical Development

I then rediscovered a field called attribution theory and that the seminal figure was Fritz Heider. I had skimmed Heider's (1958) book while a graduate student at the University of Michigan but did not take his "common sense" approach to motivation seriously. After all, Atkinson was studying Newton and seeking a mathematical representation for the universal laws of motivated behavior, whereas Heider was citing Shakespeare and Ibsen and relying on the everyday vocabulary of laypersons.

But reading Heider with the background knowledge of Rotter's research and my thoughts separating locus from control and ability from effort now placed his ideas within a more compatible framework. In Heider's (1958) naïve analysis of action, outcomes are ascribed to Can \times Try. Can, in turn, captures the relation of ability to the difficulty of the task. Thus, Heider specified three determinants of behavior: ability, task difficulty, and effort. Two of these (ability and effort) he considered internal to the person, whereas task difficulty is an external cause of an outcome.

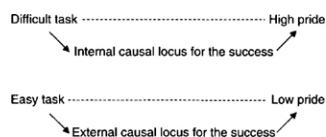
Heider and Rotter did not cite one another, although both were concerned with the perceived causes of success and failure. Rotter acknowledged one internal and one external cause, respectively skill (ability) and luck (chance), whereas Heider intuited three causes (ability, effort, and task difficulty). I combined these two lists and hypothesized four main perceived causes of achievement outcomes – ability, effort, task difficulty, and luck (see Weiner et al., 1971) Two of these are internal to the person (ability and effort) and two are external (task difficulty and luck).

Within Rotter's Expectancy/Value framework, locus of control is related to expectancy, with greater expectancy shifts (increments after success, decrements following failure) anticipated given internal rather than external control beliefs. For example, following a [p. 139 ↓] win at a tennis match he anticipated greater increments in expectancy of success than following a win at a coin toss, and similarly greater decrements after failure were hypothesized given lack of skill rather than bad luck ascriptions. Yet, closer examination of these predicted associations raised doubts. For example, if an exam is failed because the teacher is extremely harsh, it seems reasonable to hypothesize that expectancies will fall although the cause is external to the student, just as they had been hypothesized to drop when failure is attributed to the internal cause of low aptitude. Furthermore, if failure is ascribed to lack of studying and the student wants to do better in the future, then expectancy may not drop even though this cause is internal, just as expectancy had been anticipated to remain relatively unchanged if failure is attributed to the external cause of bad luck. That is, locus of control is not systematically related to expectancy and expectancy shifts because some internal causes of failure lead to large expectancy shifts (e.g., low ability) whereas others do not (e.g., lack of effort), just as some external ascriptions for failure produce major expectancy decrements (e.g., a harsh teacher) while others have little affect on subjective likelihood (e.g., bad luck).

Could there be another property or characteristic of causes rather than locus that accounts for expectancy shifts? I argued that this property is causal stability (see Weiner et al., 1971). If a cause is subject to future change, such as lack of effort (internal to the person) or bad luck (external causality), then failure would not produce downward shifts in expectancy – hope could be maintained. On the other hand, if the cause of failure is regarded as unchanging or stable, such as aptitude (internal) or a harsh teacher (external causality), then there would be an expectation of future failures and a state of hopelessness. Causal stability, I concluded, rather than causal locus, is the basis of expectancy shifts. Simply put, if the cause will prevail in the future, then the prior effect will be anticipated to recur, regardless of causal locus. I later confirmed this in experimental investigations (see Weiner et al., 1976). In this manner I bridged the past to the present and future, as had Thorndike, but with a cognitive rather than a behavioral variable.

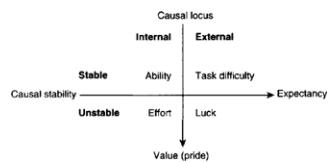
But then what is the function or consequence of causal locus? Here I returned to Expectancy/Value theory. If expectancy is accounted for by causal stability, then only the value (incentive, affect) component of Expectancy/Value theory remained, which could then be related to causal locus. Guided by Atkinson's analysis of pride, I reasoned that attributions of success to internal factors give rise to more pride than do external ascriptions – one feels greater pride after success at a tennis match attributed to high ability or high effort than following success at a match determined by chance or by the prior quality of the opponent (task ease). Hence, perception of locus of causality was the mediator accounting for Atkinson's assumption and the supporting data that pride is related to task difficulty. The harder the task, the more likely success is ascribed to the self (rather than to the ease of the task) and thus the greater the pride in accomplishment ([Figure 7.1](#)).

Figure 7.1 The harder the task, the more likely success is ascribed to the self (rather than to the ease of the task) and thus the greater the pride in accomplishment



This line of reasoning resulted in [Figure 7.2](#) (see Weiner et al., 1971), which is a 2×2 representation including four determinants of behavioral outcomes (ability, effort, task difficulty, and luck), their two properties or dimensions (causal stability and causal locus), and the linkages of expectancy to [p. 140 ↓] causal stability and value to causal locus. [Figure 7.2](#) reveals that when Rotter compared expectancy changes following success or failure given ability (internal and stable) versus luck (external and unstable) ascriptions, two dimensions or properties of phenomenal causality (locus and stability) were confounded. Rotter had incorrectly related expectancy to causal locus rather than to causal stability. The simple structure shown in [Figure 7.2](#) provided the guide for my causal thinking and was the foundation for my subsequent theory building.

Figure 7.2 A 2×2 representation of perceived causality and linked consequences



A more complete theory, I believed, needed to specify the antecedents that influence causal beliefs; all the perceived causes; additional properties or dimensions of causes; the connections between these dimensions and expectancy, affect, and other determinants of motivated behavior; and then action itself, operationalized with the usual motivational indicators of choice, intensity, and persistence of behavior. I was beginning to develop a structure that transcended content domains, included interrelated component parts, and provided a temporal sequence for a motivational episode, while retaining Expectancy/Value theory (see Weiner, 1985, 1986). These were the goals of the so-called “grand” theorists. My immediate aim was to progress in this direction.

Theoretical Elaboration

I proceeded to amend the structure in [Figure 7.2](#) by “filling in the blanks” mentioned above. Perhaps the most important addition was to include a third causal dimension, causal control, guided by my prior reasoning that locus and control must be distinguished. The control construct created some difficulties because it apparently is not orthogonal to locus. An external cause is by definition not controllable by the actor, whereas some internal causes are controllable (e.g., effort) whereas others are not (e.g., aptitude). To address the orthogonality problem, I regarded some external causes as also controllable, but by others. For example, failure due to teacher bias is external to the pupil and uncontrollable by him or her, but bias is perceived by the student as controllable by the teacher. Hence, by shifting focus, a cause can be external yet controllable. Chance, on the other hand, is an external cause not controllable by anyone.

All causes, then, are locatable within a three-dimensional taxonomic space. Considering the four main determinants of achievement outcomes: aptitude is internal, stable, and uncontrollable; effort is internal, unstable, and controllable; task difficulty generally

is regarded external, stable, and controllable (by the teacher); and luck is external, unstable, and uncontrollable. Dimensions provide the meaning or connotation of the cause. For example, aptitude “is” a property or characteristic that is internal to the person and stable that cannot be volition-ally altered. But this is a phenomenological system, so that dimensional placement depends on “how it seems to me.” Thus, if success is attributed to being a “lucky person,” then luck is classified internal, stable, and uncontrollable, thereby having the same genotypic properties as aptitude. Similarly, if effort is expended by an industrious person, then effort might be considered not only internal and controllable but also stable rather than unstable.

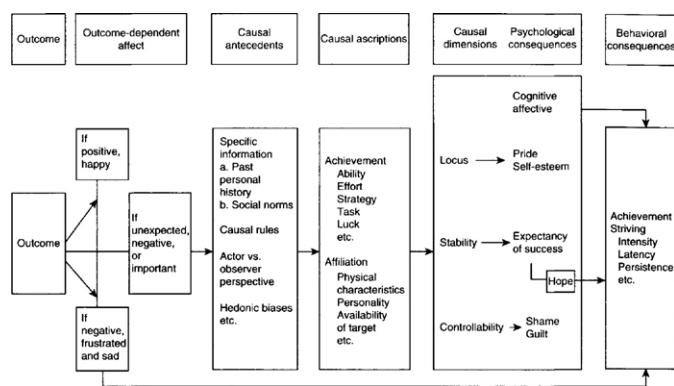
A second important addition was to expand the list of emotions linked with causal beliefs. I became acquainted with appraisal approaches to emotion and adopted the position that feelings are directed by thoughts (see Weiner et al., 1978, 1979). Thus, for [p. 141 ↓] example: attributions of success to the self generate pride; internal controllable causes for failure (e.g., lack of effort) give rise to guilt and regret; internal uncontrollable ascriptions for failure (e.g., lack of aptitude) produce shame and humiliation; stable causes of failure give rise to hopelessness; success due to an external cause controllable by others evokes gratitude; failure ascribed to an external cause controllable by others elicits anger; luck as the cause of either success or failure generates surprise; and so on (see Weiner, 2007). That is, I realized that a great deal of emotional life is guided by beliefs about causality.

In addition, it appears that some emotions, specifically happiness and unhappiness, are not linked to attributions but are tied to task outcomes. These are outcome-dependent, attribution-independent feelings. Hence, one is happy as well as surprised when succeeding because of good luck, while pride is not experienced nor is self-esteem enhanced because the cause is external to the actor. In a similar manner, failure to reach a desired goal attributed to lack of aptitude is hypothesized to produce unhappiness (outcome-dependent), a lowering of self-regard (locus-related), and shame (a consequence of internal, uncontrollable causality), along with a low expectation of future success and hopelessness and/or helplessness (stability-linked). Thus, two classes of emotion, attribution-independent and attribution-dependent, are differentiated on the basis of their cognitive antecedents (task outcome versus causal attribution for the outcome), and multiple emotions are hypothesized to coexist.

Other theory-pertinent facts meanwhile were accumulating from many researchers interested in attribution processes. Foremost among the contributors was Harold Kelley (1967), who formulated the well-known “Kelley cube.” His analysis identified social norms and past personal history among the key antecedents of causal beliefs. Kelley (1972) also examined causal schemata, or rules relating causes to effects, and these became recognized among the important determinants of causal conclusions. It also became apparent that causal understanding is not sought in all instances and is most fostered by the unexpected nonattainment of an important goal (see Gendolla and Koller, 2001).

These additions were incorporated into the formulation of an attribution-based theory of personal motivation, shown in [Figure 7.3](#). The temporal sequence in a motivated episode can be illustrated as follows. Assume a student fails an important exam. The initial experience following the failure is unhappiness. Assuming the outcome is negative and/or unexpected, there is a search for causality. Presume this person failed in the past even though she studied many hours, while others succeeded on this exam. This pattern of information gives rise to the belief that the current failure is due to lack of aptitude. Aptitude is an internal, stable, uncontrollable cause, so there is a lowering of self-esteem, low expectancy of future success, hopelessness and helplessness, and shame and humiliation. Low expectancy accompanied by these negative affects is predicted to promote the decision to drop out of school.

Figure 7.3 An attributional theory of intrapersonal behavior



Conversely, imagine that another fails the same task. This person also initially experiences unhappiness. But she has been successful in the past and laments that the night before the exam she was partying rather than studying. Hence, her current failure is ascribed to insufficient effort. This internal ascription to an unstable, controllable cause lowers personal regard but also gives rise to the maintenance of expectancy, hope, guilt, and regret, all of which are positive motivators. Hence, motivation is anticipated to increase and she is predicted to try harder in the future.

This same line of reasoning can be applied to affiliative contexts as well. For example, assume Bill calls Jane for a date and is rejected. This causes unhappiness and elicits attributional search; Bill asks: "Why won't you go out with me?" Jane may be unlikely to tell some truths because she does not want to lower Bill's self-regard by providing [p. 142 ↓] a cause internal to him such as: "You are boring" (see Folkes, 1982). But in spite of her replying: "I have to study" (an excuse or causal substitution), Bill ruminates that he has been refused many times, whereas others have girlfriends. He concludes: "I am a boring person." This ascription has the same genotypic or conceptual properties as does failing math because of perceived low ability. That is, the cause is internal (lowering esteem), uncontrollable (raising humiliation), and stable (lowering expectancies and hope). Hence, he does not seek further dates. The theory is therefore applicable across motivational domains, primarily because there is genotypic representation of causal characteristics rather than a phenotypic description of causes. That is, low math aptitude and being a boring person are conceptually identical.

Empirical Validation and Evaluative Conclusion

Does this theoretical structure receive empirical support? There is little doubt that positive and negative outcomes respectively give rise to general positive and negative emotions; that undesired and/or unexpected events promote causal search; that the cause selected is dependent on a variety of antecedents, including social norms, past history, schematic beliefs, and hedonic biases; and that locus, stability, and control are key causal properties (although there may be other causal characteristics as well, such as globality or causal generality across situations). Causal locus relates to pride

and self-esteem; causal stability in part determines expectancy shifts and the extent of hope, hopelessness, and helplessness; and causal control is linked with affects including shame (to uncontrollable causality), as well as guilt, and regret (to controllable causality). There are voluminous literatures supporting these assertions (see Weiner, 1985, 1986, 1995).

However, in spite of this array of evidence and theoretical generality, I would be remiss, perhaps even intellectually dishonest, if I did not share a misgiving. I am very concerned about the last links in the theory going from expectancy and affect to performance. Perhaps I am overly worried about these [p. 143 ↓] associations given the secure place of Expectancy/Value theory in motivation history.

What is unclear are the associations of expectancy and attribution-linked emotions to achievement behaviors such as academic performance and dropping out of school. With regard to expectancy, motivational theorists have offered contradictory predictions. It may be that low expectancy of success or a difficult task generates most effort (see Locke and Latham, 1990); or, as Atkinson (1957) intuited, perhaps motivation is greatest when tasks are of intermediate difficulty; or perhaps being close to a goal (high expectancy) enhances performance (Lewin, 1938). In sum, the relation between expectancy and intensity of performance and choice behavior is uncertain. In regards to affect, similar difficulties arise. It is possible that high guilt impedes rather than facilitates performance, or that high shame is a positive motivator in certain settings, or even that shame and guilt cannot be readily distinguished inasmuch as they correlate quite highly in experimental research. Few studies have tested the entire theory with appropriate path analytic techniques (see Weiner, 1986), which in part accounts for my fear that the whole theory may be less than the sum of its parts because of the weak final links to motivated behavior.

There is another theoretical message in the massive support for the component associations in the theory accompanied by the uncertain validity of the entire conception. That is, it is not possible to construct a complete theory of achievement motivation. Achievement striving is impacted by causal beliefs but also is determined by other factors including anticipated costs and benefits (e.g., a promise of increased allowance from a parent could raise grades); affiliative concerns (e.g., studying with a potential mate may increase library time); the need to earn money (which

impedes performance by decreasing the time available for study); and on and on. Overdetermination renders it exceedingly difficult to significantly relate causal beliefs to molar achievement indices such as grade point average and dropping out of school.

What are the best predictors of school dropout and other signs of achievement striving in school settings? My guess is that these variables are school identification and feelings of belonging, peer-group norms, parental guidance and goals, and other determinants tied to societal values and socioeconomic class. This is not to imply that attributions have no effects. After all, attribution change programs have even proven effective in altering the performance of some students, as discussed later in this paper. In addition, attributions do explain aspects of emotional life in the classroom and expectations of success – no small feats.

This same general analysis is applicable to other motivation domains as well. Hence, the implicit position of this theoretical approach is that one set of principles (attribution-consequence relations) has been isolated that contribute to successful prediction in a wide variety of settings, yet these associations often are not the sole determinants of the behavior in question. Many principles stand side by side that are pertinent to the understanding of achievement strivings, which is why attribution thinking alone cannot be predictive of (or will only weakly predict) school performance. This conclusion must be taken into account by all motivation-based theories of school performance.

A Theoretical Turning Point

I now shift focus and offer a different, albeit related, theoretical development. As just indicated, I was disturbed by the difficulty in predicting school-related achievement performance and about the possibility that most variance in predictions of school motivation and dropout might be accounted for by broad demographic, socioeconomic indices rather than by hypothesized psychological mediators. I thus explored opportunities to shift [p. 144 ↓] research from achievement to other human arenas. In doing so, I drifted further and further away from my historical roots and identification as an achievement theorist, thereby making comparisons with Atkinson's theory more problematic.

When thinking about future theoretical and empirical directions, I recognized that I was approaching achievement motivation from an intrapersonal perspective. The pertinent thoughts and emotions did not presume the presence or influence of others (with small exceptions such as the use of social norms as an antecedent of causal conclusions or social comparisons affecting emotional experiences such as shame). A research participant could be tested in an isolated room, with success and/or failure manipulated, and it was believed that attributions for these outcomes, as well as linked expectancies and emotions, would be generated that influence subsequent achievement striving. This is an asocial theory.

Yet in my first extended attribution research discussed earlier (Weiner and Kukla, 1970), participants assuming they were teachers evaluated students following hypothetical success or failure ascribed to various combinations of ability and effort. This judgment research concerned reactions toward others. Although the findings were essential to the generation of an intrapersonal theory, they in fact had little to do with personal achievement strivings but related to motivated behavior toward others, or social motivation. I had been thinking of asocial motivation but at times engaged in social motivation research; that is, I failed to recognize an intrapersonal/interpersonal distinction in the study of motivated behavior. My subsequent theory building explicitly turned from intrapersonal to interpersonal behavior and from achievement strivings to social motivation, particularly to help-giving (which I discuss here) and, to a lesser extent, aggression, impression management strategies such as giving excuses, and punishment (which I do not discuss here because of space limitations; see Weiner, 1995, 2006). I hoped that findings in these areas would not only verify the associations between the already discovered components within the theory but also would provide support for the entire conception. And this proved to be the case!

The Empirical Departing Point to Social Motivation

The initial studies explicitly in the domain of social motivation that I conducted returned to the ability/effort distinction explored by Weiner and Kukla (1970) and the earlier research by Schmitt (1964). It was evident from these publications that lack of effort as

the cause of a personal or social failure (doing poorly on an exam; nonrepayment of a debt) results in greater social disapproval than achievement or social failure because of broadly defined lack of ability (e.g., low school capability; no money because of illness). Viewing these findings from the intrapersonal theory perspective shown in [Figure 7.3](#), it appeared that ability and effort give rise to different evaluations because of their contrasting placements on the causal dimension of control. Lack of effort is a controllable cause; it “could have been otherwise.” On the other hand, the absence of aptitude (or illness) is uncontrollable and not subject to volitional change. Thus, the same theoretical analysis is applicable in these two diverse contexts (see Weiner, 1995, 2006), again because phenotypically distinct causes share genotypic representations regarding causal controllability. But remember that in this case the behavior of the other, and not the self, is being considered. Hence, self-directed affects such as guilt and shame are inappropriate to include. How, then, should emotions be represented in an interpersonal context?

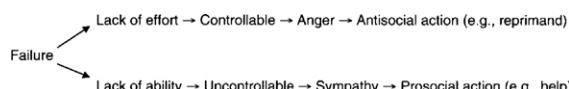
Incorporating Emotions

The emotion appraisal literature provides extensive support for the hypothesis that [p. 145 ↓] control beliefs regarding others have emotional consequences. Specifically, if another commits a transgression and the cause is under personal control, whether school failure due to lack of effort or coming late for an appointment because it was forgotten, then involved observers react with anger. One is angry when one's child fails at school because of not studying! On the other hand, if the cause of a transgression was not subject to volitional change, whether it is school failure attributed to low aptitude or being tardy for an appointment because the bus broke down, then observers react with sympathy (see, for example, Weiner et al., 1982). Consider your own reactions to the academic failure of a child with a mental handicap. A more complete social motivation analysis incorporating emotional reactions to causal thoughts regarding control by others is represented as follows, as in the intrapersonal theory:

Event → Cause → Causal controllability →
Emotion → Action

More specifically, considering observer reactions to the achievement striving of others, this sequence can be depicted as shown in [Figure 7.4](#).

Figure 7.4 Social motivation analysis incorporating emotional reactions to causal thoughts regarding control by others, considering observer reactions to the achievement striving of others



Refining Controllability

The concept of causal control appears similar to personal responsibility. In earlier paragraphs the phrase “could have been otherwise” was used to elucidate or define this construct. One is responsible (able-to-respond) for not expending effort, but not for lacking aptitude, which cannot be changed merely by willpower. On the other hand, the philosophical literature and writings on criminal justice clearly point out that control is distinct from responsibility. This is in part because there are controllable actions for which the individual is either not held responsible or responsibility is diminished. For example, responsibility is negated or weakened if a social transgression is in service of a higher moral goal (labeled a justification). One is justified in coming late for an appointment if a roommate became ill and had to be driven to the hospital. In this case, coming late is not regarded a “moral failure.” In a similar manner, if a crime is committed because the individual is mentally unstable, or does not understand the difference between right and wrong, then again the act may be controllable although the person is perceived as not responsible (see Weiner, 1995, 2006).

This reasoning promoted the inclusion of an additional path within the theory, one connecting control to responsibility. The control-responsibility separation provided the opportunity to incorporate moderators of this relation, even though on most occasions control and responsibility have the same value and meaning. The elaborated theory, including only the causal dimension of control, is depicted as:

Event → Causal antecedents → Cause → Causal dimension (control) → Personal responsibility → Emotion → Action

[p. 146 ↓]

Testing the Interpersonal Theory

The testing of the interpersonal theory is next documented in some detail to present a contrast to the discussion of the intrapersonal theory. Research attention was first given to help-giving because altruism and prosocial behavior form the core of social motivation. Two specific research contexts were examined: (1) help-giving in achievement situations (e.g., lending class notes); and (2) charity aid to the needy (e.g., financial help to various stigmatized individuals).

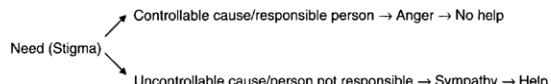
The logic of the experimental manipulations in these two settings was identical. Information was varied that created disparate beliefs about causal control/personal responsibility. In the achievement-related studies, pupils were portrayed as asking to borrow class notes either because they “went to the beach” (which elicits beliefs in personal control and responsibility) or because of eye problems (which gives rise to attributions of uncontrollability and inferences of nonresponsibility; see Schmidt and Weiner, 1988; Weiner, 1980). In the charity research, reactions toward individuals with stigmas that are apparently produced by controllable or volitionally undertaken behaviors (e.g., alcoholism, obesity) were compared with reactions to stigmas associated with uncontrollable states or conditions (e.g. blindness, Alzheimer's disease). An alternative to offering contrasting stigmas was to present research participants with identical stigmas accompanied by different causes of the conditions (e.g., obesity because of overeating versus obesity caused by a thyroid problem; AIDS due to promiscuous sexual behavior versus AIDS caused by a transfusion with contaminated blood; see Weiner et al., 1988).

Research participants then rated the controllability of and/or personal responsibility for the cause of the need for class notes or the cause of the stigma, their affective reactions of anger and sympathy, and their prosocial tendencies (likelihood of lending the notes; amount of charity allocation). Because hypothetical situations were created,

behavioral determinants such as real time available to help and the actual amount of money that one has were not anticipated to affect the judgments. That is, other determinants of “behavior” were rendered irrelevant, thus overcoming some problems that plagued research regarding achievement strivings. Indeed, the studies in which I was involved often used simulation procedures – research participants “pretended” to be enrolled in the same class as the student in need or “pretended” to be members of a charity board dispensing funds. This simulation procedure raises hackles among some psychologists for both valid and invalid reasons that cannot be addressed here. Thankfully, others engaging in similar pursuits performed “real” manipulations (assuming that an experimental intervention can be labeled “real”), while some studies involved observations without personal intervention.

In contrast to the paucity of research investigations examining the full intrapersonal theory, many scientists conducted help-related research that assessed attributions as well as emotions and/or behaviors. Although the reported research studies did not manipulate or measure all the components in the hypothesized theory, the most essential mediational concepts were taken into account. The motivation sequences typically tested are shown in [Figure 7.5](#).

Figure 7.5 Typically tested motivational sequences



Along with three colleagues (Rudolph et al., 2004), we conducted a meta-analysis of these research studies. To be included in the analysis, the investigation had to have at least one attribution variable, such as controllability; at least one emotion synonymous with anger or sympathy; and a relevant behavior. Also, first-order correlations were required. Our search found 39 usable studies, including nearly 8,000 research participants. A number of path models were examined, two of which are presented here – Model 1, which has direct links between thinking as well as affect and doing; and Model 2, [p. 147 ↓] which has no thinking-action path so that affect is the only proximal determinant of behavior. Thus, the two models differ in whether thinking directly or only indirectly affect action.

Table 7.1 shows the path coefficients between the attribution-affect-behavior variables, all of which are significant. The table reveals that controllability is negatively related to sympathy ($\beta = -0.45$) and positively to anger ($\beta = 0.52$) for both models. That is, the more controllable the need or stigma (e.g., poverty because of laziness as opposed to a lack of available jobs, or AIDS because of promiscuous sexual behavior rather than a blood transfusion), the less the sympathy and the more the anger. In addition, sympathy positively relates to help-giving (average $\beta = 0.38$), while anger relates negatively to help (average $\beta = -0.08$). **Table 7.1** also shows that the path between thoughts about control and behavior is relatively weak ($\beta = -0.05$); including this linkage does not improve the fit of the models.

	Model 1	Model 2
Control-Sympathy	-0.45	-0.45
Control-Anger	0.52	0.52
Control-Help-giving	-0.05	–
Sympathy-Help-giving	0.37	0.39
Anger-Help-giving	-0.07	-0.09

In short, the full pattern of data is very consistent with the theory and, guided by the rule of parsimony, supports a thinking → feeling → acting motivation sequence – thoughts direct feelings and feelings guide action. Thoughts do play a role in helping behavior, but only as distal determinants through their influence on emotion. I regard the thought (attribution) → affect → action sequence as the “deep structure,” or the basic genotypic representation, of a motivation episode. However, I certainly recognize that not all emotions require causal antecedents and that affect may influence thinking (i.e., there is an affect-cognition sequence). These beliefs nonetheless are relatively peripheral to this theoretical approach.

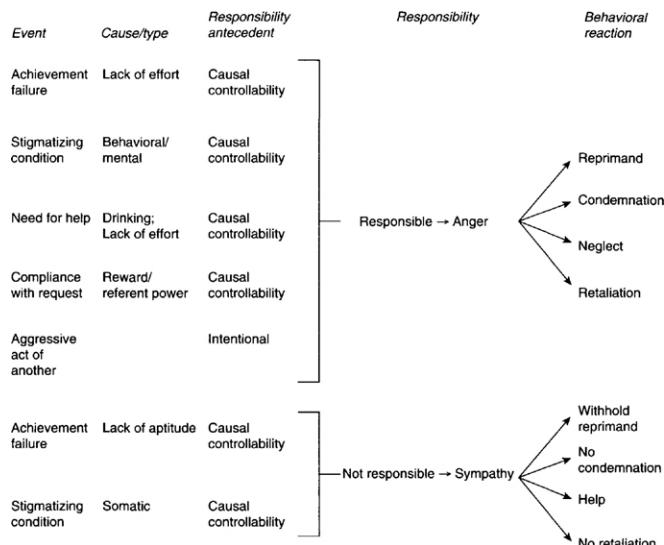
In related social motivation research examining aggression, the main findings reported in the studies of help-giving were replicated. However, a meta-analysis of aggression studies also found a direct link between attributions and action (see Rudolph et al., 2004; Weiner, 2006). Thus, although the basic structure of a motivation sequence is

captured by the theory, much remains to be determined regarding the dynamics of behavior in specific motivation settings.

The Complete Interpersonal Theory

What, then, is the structure of the interpersonal theory and the contexts in which it has been validated? In addition to achievement judgments, help-giving, and aggression, the interpersonal theory has been extended to an analysis of the basis of social power (Rodrigues, 1995; Rodrigues and Lloyd, 1998) and to broader examinations of stigmas, although these areas of research are not reviewed here (see Weiner, 1995, 2006). They are, however, included in the complete [p. 148 ↓] theory, which is shown in [Figure 7.6](#). The upper portion of [Figure 7.6](#) portrays situations in which the actor is personally responsible for a negative outcome. In the first row, the domain is achievement and the failed outcome is ascribed by an observer to lack of effort. Effort is a controllable cause and, in the absence of mitigating information, the person is held responsible for the failure. If there also is some personal involvement (e.g., the failing individual is your child), then the reaction elicited by low effort is anger. Anger, in turn, promotes an antisocial reaction such as punishment or rejection.

Figure 7.6 An attributional theory of interpersonal behavior



In contrast, the bottom portion of Figure 7.6 portrays instances in which the actor is not held responsible for the negative outcome. In achievement settings, the uncontrollable cause of failure is lack of aptitude. Since aptitude is not controllable by the failing person, that individual is not held personally responsible for the negative outcome. Lack of responsibility for failure gives rise to sympathy and pity. These emotions, in turn, promote prosocial responding such as help-giving and comfort. This same general analysis applies to the other content domains included in Figure 7.6, where the logic of the theory is extended to explain help-giving, aggression, and social power.

In sum, this theory has interrelated component parts, generalizes to wide-ranging motivational contexts, and receives extensive empirical support. At the most molar level, the theory states that motivated behavior is a function of cognition and affect. This perhaps provides one among the many alternatives to the equally broad Lewinian statement [p. 149 ↓] that behavior is a function of the person and the environment. A somewhat more specific formulation of this general law is that behavior is determined by causal reasoning and responsibility inferences, along with their linked emotions. In addition, emotions are proximal in accounting for action, while the proximal versus distal role of thinking remains to be determined and appears to differ in altruistic versus aggressive contexts.

Comparing and Contrasting the Intrapersonal and Interpersonal Theories

The intrapersonal and interpersonal attribution theories of motivation portrayed respectively in [Figures 7.3](#) and [7.6](#) have much in common. For both, a motivational episode is initiated with a prior outcome; that is, the analysis is postperformance yet prior to the next action. Given an unexpected, negative, and/or important event, there is a search for causality. The cause selected is then analyzed according to its causal properties. This categorization gives rise to emotions, which are proximally linked to action. Hence, the structures of the intrapersonal and interpersonal theories are virtually identical and, for both, causal beliefs provide the foundation for theoretical analysis.

On the other hand, the two theories have important theoretical and empirical differences. At the conceptual level, the intrapersonal theory includes outcome-dependent affects not necessarily part of the interpersonal conception; the information used to make causal inferences about others differs from information used regarding the self inasmuch as personal causal antecedents are more readily available and there is the possibility of hedonic biasing; in the interpersonal theory the dimension of importance is controllability, while locus and stability are also included in the intrapersonal theory; because stability and expectancy are not part of the interpersonal theory, the Expectancy/Value framework is maintained only in the intrapersonal conception (although these other dimensions could be incorporated within the interpersonal theory); in the interpersonal but not the intrapersonal conception, there is an added step linking control to responsibility, permitting the inclusion of mitigating factors; and given the intrapersonal theory, the emotions elicited are directed toward the self, whereas in the interpersonal theory the emotions involve others.

In addition to these theoretical contrasts, the theories differ in their empirical support. The intrapersonal approach has been primarily tested in achievement contexts, whereas the interpersonal framework has been empirically examined in a wide variety of motivational domains including help-giving, aggression, and social power. Because the interpersonal theory often is concerned with the judgments of others, the theory is more readily testable and has found greater support than the intrapersonal theory,

which has attempted to predict the more elusive achievement performance. Thus, for the interpersonal theory, the whole indeed is more than the sum of the parts, whereas this may not be the case given the intrapersonal perspective. But both provide rich and complementary conceptual frameworks for the study of human motivation. They qualify for inclusion among the grandiose attempts to provide a unifying perspective to understand motivated behavior.

Theoretical Applications

My motivational hero, Kurt Lewin, stated somewhere that “there is nothing as practical as a good theory.” It is one of the few things he wrote that I do not believe. In fact, there is nothing as practical as a simple engineering principle, without theoretical burden. For example, applications of psychology have perhaps been most fostered by Skinnerian rules regarding the consequences of reward [p. 150 ↓] and punishment. Getting a pigeon (or a human) to respond by reinforcing a response with a desirable incentive is a very effective method for behavior change. It is a proven technique in many clinical studies and has been used to investigate a variety of problems. Yet it is known that Skinner hated theories. On the other hand, there have been few (if any) real-life applications or consequences resulting from Drive or Expectancy/Value theory, and the same lack of applied value can be concluded about Lewin's formal theory of motivation.

I must admit that as a motivational theorist, I have had little interest in practical application. I played a game of chess and struggled with the creation of a viable theory – parsimonious, generalizable, empirically supported. Nevertheless, it is very satisfying when attribution theory is successfully applied to a problem area, particularly when the theoretical extension had not been foreseen.

A number of applications are ongoing and promising. It must be said, however, that the utilization potential far exceeds the applications in progress. In the current theoretically originated programs, the focus is on altering dysfunctional causal attributions to those that are more adaptive. The underlying principle guiding utilization of the theory is that if there is a change in perceived causality, then there also will be a change in subsequent behavior. This rule has been applied to achievement strivings; health maintenance

among the elderly; treatment of the mentally ill and other stigmatized groups; help-giving; aggression; and even to behaviors in the airport, the hospital, and the classroom, thereby capturing the vast range of attribution theory.

Corresponding to the intra- and interpersonal theories, the applications will be grouped according to their goal of changing self-perception versus altering ascriptions about others. There is a wonderful paradox in the goals of these programs; when altering self-perceptions, the typical aim is to shift the cause of nonattainment of a goal so it no longer is perceived as uncontrollable. For example, in achievement settings, the aim is to shift attributions for failure from lack of ability to lack of effort. On the other hand, when changing other perceptions, the desire often is to have the cause no longer be perceived as controllable. For example, when considering stigmas the aim is to reduce the blaming of the victim.

Self-Perception: Pupils regarding Classroom Performance and the Elderly regarding Health

There are demotivating (dysfunctional) and motivating (adaptive) causal beliefs, particularly in regard to achievement failure. As already extensively discussed, perhaps the least adaptive ascription to failure is lack of aptitude. On the other hand, lack of effort as the inferred cause of failure is perhaps most adaptive. The contrasting consequences of ability versus effort ascriptions for failure have given rise to a variety of attribution change programs that attempt to shift attributions for failure from low ability to insufficient effort; that is, from stable and uncontrollable to unstable and controllable.

These programs have varied formats, but the technique with the widest usage shows video clips of students discussing their prior failures and communicating the realization that their failure was not due to low ability but rather was caused by insufficient effort or poor strategy. The filmed students also state that this causal shift greatly improved their school performance. Perry et al. (1993) document that this intervention, along with related information, has lasting positive consequences on grade point average in a

college setting. This effect has been found in a large number of research studies (see Van Overwalle et al., 1989).

In contrast to shifting causal beliefs to the internal, controllable causes of effort and strategy, some treatments have followed a different path, attempting to alter attributions [p. 151 ↓] for poor performance among freshmen to the initially harsh grading policies of the college, which are described as becoming more lenient over school years. That is, manipulating the perceived cause of failure from stable to unstable is the goal of the experimental treatment, which theoretically increases expectancy of success and hope without asking the students to alter their behavior. Positive treatment effects also have been reported from this intervention (see Wilson et al., 2002).

Health Maintenance

Achievement is typically associated with school or job performance. Yet success and failure have a variety of indicators and meanings. Another application of attribution change programs targets health maintenance and walking among the elderly. It is known that exercise has great benefits for all, and particularly for the elderly. Yet the aged often do not engage in physical activities, attributing their “failure” (inactivity) to old age, which is seen as internal, stable, and uncontrollable. This attribution has the same conceptual properties and dysfunctional consequences as an attribution for failure in the classroom to lack of aptitude. Programs thus have been initiated to persuade the elderly to ascribe their inactivity to insufficient effort or “not trying” rather than to their age. A variety of techniques are used to accomplish this change, particularly via information directly transmitted by health professionals (see Sarkisian et al., 2007). This intervention has proven effective in increasing exercise and walking, extremely important changes for health maintenance. However, the research supporting this claim is yet minimal.

In sum, attribution change programs enhance student performance in the classroom and perhaps elderly “performance” regarding exercise. The attribution intervention techniques are simple, require little cost, and promise great benefit. This is an important direction of application that needs more attention.

Teacher-Training Programs

It has been documented that some prevalent teacher practices result in students making low ability attributions for their poor performance, the most harmful of the attributions for failure (see Graham, 1990). Such practices include not punishing and conveying sympathy given failure at an easy task, overly praising success at an easy task, and providing help when it is not sought. When reprimand for failure at an easy task is withheld, the attribution for the poor performance is perceived as something other than a lack of effort. Furthermore, the conveyance of sympathy provides a cue that the failure was uncontrollable. Given these communications, the student is likely to conclude that the teacher believes his or her failure was caused by an absence of ability. This causal communication then increases the pupil's own belief that he or she is unable. In a similar manner, pre-emptive help-giving is a cue that the pupil "cannot," while praise for success at an easy task conveys the cause of this outcome was high effort (which implicates a low level of ability).

In sum, there are a variety of apparent teacher practices in need of alteration from an attribution perspective. Communication of low ability is not the intent of these teacher behaviors and the actions appear "kind" (e.g., withholding blame, giving praise, and providing unsought help). These principles and other extrapolations from attribution theory presently are being incorporated into some teacher training programs to hopefully alter teacher behavior and, in turn, the potentially harmful low ability self-perceptions of their pupils. But the effects of this inclusion remain to be systematically examined and the research is scanty.

Other Perception

As already noted, rather than altering causal ascriptions from uncontrollable and stable (ability) to controllable and unstable (effort), thereby affecting expectancy and affect, the [p. 152 ↓] goals of programs regarding other perception often are to shift attributions (e.g., for illness) from controllable (e.g., lifestyle) to uncontrollable (e.g., genetics). Theoretically, this change reduces perceptions of responsibility and anger, increases

sympathy, and eliminates observer antisocial behavior (or promotes prosocial behavior). Many populations have been the targets of this type of attribution change in research studies. Again, however, there has not been a sufficient number of programs that apply this knowledge to reach any scientific closure. The rather sparse change attempts are the focus of the following discussion.

The Mentally Ill

In the current climate of the growth of neuropsychology, biological models of major mental disorders such as schizophrenia and depression predominate, along with drug treatments. Yet the majority of the public may still perceive the mentally ill as morally weak and responsible for their problems (see Neff and Husaini, 1985). This is reflected in critical and hostile comments (labeled expressed emotion, or EE) directed by family members toward the victims of these mental disorders. Furthermore, it is well established that "living in a high EE-home environment more than doubles the baseline relapse rate for schizophrenic patients" (Barrowclough and Hooley, 2003: 849).

Inasmuch as EE is elicited by controllable beliefs regarding the illness of others and harms patients by promoting relapse, it then follows that programs altering ascriptions of mental illness from controllable to uncontrollable should reduce recidivism. From the viewpoint of attribution theory, this intervention reduces inferred responsibility and the expression of anger, while increasing sympathy, warmth, and prosocial actions. Some interventions have been undertaken to convince caretakers of the uncontrollability of the problems of their mentally ill family members (see Medvene and Krauss, 1989), and it has been reported that decreasing EE does lessen relapse (see Pitschel-Walz et al., 2001). What is needed is a rash of such intervention programs.

Other Stigmas

Those with major mental disorders are not the only stigmatized individuals who additionally suffer because they are regarded as moral failures. Across a variety of stigmas, the most defining characteristic that promotes rejection and social distancing is the perception of personal responsibility (Feldman and Crandall, 2007). Many illnesses

are believed to be caused by lifestyle choices that are amenable to personal control. For example, those with lung cancer are blamed because of the association of this disease to smoking (see Cooper, 1984; Kim and Shanahan, 2003), which produces family conflict. And many stigmatized are not given opportunities to gain jobs and rent apartments because they are perceived as "immoral" (see Corrigan et al., 2003). The antisocial reactions toward those perceived as responsible for their stigmas has even been observed in the actions of health professionals, including less favorable medicine-dispensing decisions among doctors (see Brewin, 1984; Brewin and Antaki, 1982) and reduced caring behavior of nurses (Harborne, 1996). Unfortunately, intervention attempts targeting these punitive attribution-driven reactions have not been reported.

Aggression

As documented earlier in this chapter, aggression is one response following an inference of other-person responsibility for a negative outcome and the anger that this elicits. Hence, one method to reduce aggression is to lessen perceptions of other-person responsibility for negatively valenced events.

In a business-related utilization of the principle, Folkes (1984) demonstrated that consumer reactions to product failure are linked with the perceived cause of their plight. Often, for example, flight delays are perceived by passengers as caused by [p. 153 ↓] company shortcomings, which generates anger and consumer complaints. Hence, company adoption of simple communication principles, such as announcing the flight delay is due to weather problems (an excuse shifting attributions from controllable to uncontrollable) will reduce anger and passenger complaints (one indicator of aggressive retaliation).

Evidence of these attribution → affect → behavior relations also is displayed in more serious aggressive acts. For example, there are some data that abusive mothers tend to overattribute controllability to their children regarding acts such as the accidental spilling of milk, which then causes anger and abusive actions (see Bradley and Peters, 1991). Treatment programs for abusive parents often focus on anger management, but an attributional perspective suggests that the cognitive antecedents of anger also should be targeted.

An overattribution of intention and responsibility for negative actions also has been documented among aggressive adolescents (see Dodge and Crick, 1990). Guided by this finding, cognitive intervention programs have been designed to reduce intentionality inferences ("he stepped on my toe on purpose") among aggressive-prone children (see Graham and Hudley, 1992). Part of the program includes training to discriminate between intentionally caused versus unintentional negative actions in situations of causal ambiguity. Such training programs do decrease subsequent aggression. In addition, Graham et al. (1995) report that aggressive children benefit from learning the adaptive value of providing excuses so that their own negative behavior is not perceived by others as controllable or intentional. As has been true throughout this discussion, it is unfortunate that the body of work targeted to reduce aggression is not sufficient to draw conclusions about the "success" of attribution interventions. On the positive side, the available research is certainly reasonable and encouraging.

Some Concluding Thoughts

Carved on the National Archives Building in Washington, DC is a saying from William Shakespeare's *The Tempest*: "What's past is prologue." For an attribution theorist, this means that the interpretation of the past, and particularly the perceived causes of prior events, determines what will be undertaken in the future. I have attempted to create conceptual structures to capture this truism. The defining characteristics of these structures include testable hypotheses, generalization to a wide range of behavioral domains and both intrapersonal and interpersonal actions, parsimony, incorporation of thoughts and emotions, and potential for application. These qualities make me happy (outcome-dependent), grateful (attribution of success to help from others), and proud (attribution of success to a great deal of effort). Thus, the theory accurately predicts the feelings and actions of at least one of its creators!

References

- Atkinson, J.W. (1957) Motivational determinants of risk-taking behavior . Psychological Review , 64, 359–372. <http://dx.doi.org/10.1037/h0043445>

- Atkinson, J.W. and Birch, D. (1970) *The Dynamics of Action*. New York: Wiley.
- Barrowclough, C. and Hooley, J.M. (2003) Attributions and expressed emotion: A review. *Clinical Psychology Review*, 23, 849–880. <http://dx.doi.org/10.1016/S0272-7358%2803%2900075-8>
- Bradley, E. and Peters, R. (1991) Physically abusive and nonabusive mothers' perceptions of parenting and child behavior. *American Journal of Orthopsychiatry*, 61, 455–460. <http://dx.doi.org/10.1037/h0079263>
- Brewin, C.R. (1984) Perceived controllability of life-events and willingness to prescribe psychotropic drugs. *British Journal of Social Psychology*, 23, 285–287. <http://dx.doi.org/10.1111/j.2044-8309.1984.tb00642.x>
- Brewin, C.R., ed. and Antaki, C. (eds) (1982) *Attributions and Psychological Change*. London: Academic Press.
- Cooper, E.T. (1984) A pilot study of the effects of the diagnosis of lung cancer on family relationships. *Cancer Nursing*, 7, 301–308. <http://dx.doi.org/10.1097/00002820-198408000-00004>
- Corrigan, P.W., Markowitz, F.E., Watson, A.C., Rowan, D. and Kubiak, M.A. (2003) An attribution model of public discriminations towards persons with mental illness. *Journal of Health and Social Behavior*, 44, 162–179. <http://dx.doi.org/10.2307/1519806>
- de Charms, R. (1968) *Personal Causation*. New York: Academic Press.
- Dodge, K.A. and Crick, N. (1990) Social information-processing biases of aggressive behavior in children. *Personality and Social Psychology Bulletin*, 16, 8–22. <http://dx.doi.org/10.1177/0146167290161002>
- Feldman, D.B. and Crandall, C.S. (2007) Dimensions of mental illness stigma: What about mental illness causes social rejection? *Journal of Social and Clinical Psychology*, 26, 137–154. <http://dx.doi.org/10.1521/jscp.2007.26.2.137>

- Folkes, V.S. (1982) Communicating the causes of social rejection . Journal of Experimental Social Psychology , 18, 235–252. <http://dx.doi.org/10.1016/0022-1031%2882%2990052-X>
- Folkes, V.S. (1984) Consumer reactions to product failure: An attributional approach . Journal of Consumer Research , 10, 398–409. <http://dx.doi.org/10.1086/208978>
- Gendolla, G.H.E. and Koller, M. (2001) Surprise and causal search: How are they affected by outcome valence and importance? Motivation and Emotion , 24, 237–250.
- Graham, S. (1990) On communicating low ability in the classroom: Bad things good teachers sometimes do . In S. Graham, ed. and V.S. Folkes (eds), Attribution Theory: Applications to Achievement, Mental Health, and Interpersonal Conflict , pp. 17–36. Hillsdale, NJ: Lawrence Erlbaum.
- Graham, S. and Hudley, C. (1992) An attributional approach to aggression in African-American children . In D. Schunk, ed. , and J. Meece (eds), Social Perceptions in the Classroom: Causes and Consequences , pp. 75–94. Hillsdale, NJ: Erlbaum Press.
- Graham, S., Weiner, B. and Benesh-Weiner, M. (1995) An attributional analysis of the development of excuse giving in aggressive and nonaggressive African American boys . Developmental Psychology , 31, 274–284. <http://dx.doi.org/10.1037/0012-1649.31.2.274>
- Harborne, A. (1996) Challenging behavior in older people: nurses' attitudes . Nursing Standard , 11, 39–43.
- Heider, F. (1958) The Psychology of Interpersonal Relationships . New York: Wiley. <http://dx.doi.org/10.1037/10628-000>
- Hull, C.L. (1943) Principles of Behavior . New York: Appleton-Century-Croft.
- Kelley, H.H. (1967) Attribution theory in social psychology . In D. Levine (ed.) Nebraska Symposium on Motivation , pp. 192–238. Lincoln: University of Nebraska Press.

Kelley, H.H. (1972) Causal schemata and the attribution process . In E.E. Jones, ed. , D.E. Kanouse, ed. , H.H. Kelley, ed. , R.E. Nisbett, ed. , S. Valins, ed. , and B. Weiner (eds), Attribution: Perceiving the Causes of Behavior , pp. 151–174. Morristown, NJ: General Learning Press.

Kim, S. and Shanahan, J. (2003) Stigmatizing smokers: Public sentiment toward cigarette smoking and its relationship to smoking behaviors . Journal of Health Communications , 8, 343–367. <http://dx.doi.org/10.1080/10810730305723>

Lewin, K. (1938) The Conceptual Representation and the Measurement of Psychological Forces . Durham, NC: Duke University Press. <http://dx.doi.org/10.1037/13613-000>

Locke, E. and Latham, G. (1990) A Theory of Goal Setting and Task Performance . Englewood Cliffs, NJ: Prentice-Hall.

Medvene, L.J. and Krauss, D.H. (1989) Causal attributions and parent-child relationships in a self-help group for families of the mentally ill . Journal of Applied Social Psychology , 19, 1413–1430. <http://dx.doi.org/10.1111/j.1559-1816.1989.tb01456.x>

Neff, J.A. and Husaini, B.A. (1985) Lay images of mental health: Social knowledge and tolerance of the mentally ill . Journal of Community Psychology , 13, 3–12. <http://dx.doi.org/10.1002/1520-6629%28198501%2913:1%3C3::AID-JCOP2290130102%3E3.0.CO;2-0>

Perry, R.P., Hechter, F.J., Menec, V.H. and Weinberg, L.E. (1993) Enhancing achievement motivation and performance in college students: An attributional retraining perspective . Research in Higher Education , 34, 687–723. <http://dx.doi.org/10.1007/BF00992156>

Pitschel-Walz, G., Leucht, S., Bauml, J., Kissling, W. and Engel, R.R. (2001) The effect of family interventions on relapse and rehospitalization in schizophrenia – meta-analysis . Schizophrenia Bulletin , 27, 73–92. <http://dx.doi.org/10.1093/oxfordjournals.schbul.a006861>

Rodrigues, A. (1995) Attribution and social influence . Journal of Applied Social Psychology , 25, 1567–1577. <http://dx.doi.org/10.1111/j.1559-1816.1995.tb02633.x>

Rodrigues, A. and Lloyd, K.L. (1998) Reexamining bases of power from an attributional perspective . Journal of Applied Social Psychology , 28, 973–997. <http://dx.doi.org/10.1111/j.1559-1816.1998.tb01663.x>

Rotter, J.B. (1954) Social Learning and Clinical Psychology . Englewood Cliffs, NJ: Prentice-Hall. <http://dx.doi.org/10.1037/10788-000>

Rotter, J.B. (1966) Generalized expectancies for internal versus external control of reinforcement . Psychological Monographs , 80, 1–28.

Rudolph, U., Roesch, S.C., Greitemeyer, T. and Weiner, B. (2004) A meta-analytic review of help giving and aggression from an attributional perspective . Cognition and Emotion , 18, 815–848. <http://dx.doi.org/10.1080/02699930341000248>

Sarkisian, C.A., Prohaska, T.R., Davis, C. and Weiner, B. (2007) Pilot test of an attributional retraining intervention to raise walking levels in sedentary older adults . Journal of the American Geriatrics Society , 55, 1842–1846. <http://dx.doi.org/10.1111/j.1532-5415.2007.01427.x>

Schmidt, G. and Weiner, B. (1988) An attribution-affect-action theory of behavior: Replications of judgments of help giving . Personality and Social Psychology Bulletin , 14, 610–621. <http://dx.doi.org/10.1177/0146167288143021>

Schmitt, D.R. (1964) The invocation of moral obligation . Sociometry , 27, 299–310. <http://dx.doi.org/10.2307/2785620>

Spence, K.W. (1956) Behavior Theory and Conditioning . New Haven, CT: Yale University Press. <http://dx.doi.org/10.1037/10029-000>

Thorndike, E.L. (1911) Animal Intelligence . New York: Macmillan.

Tolman, E.C. (1932) Purposive Behavior in Animals and Men . New York: Appleton-Century-Croft.

Van Overwalle, F., Segebarth, K. and Goldchstein, M. (1989) Improving performance of freshmen through attributional testimonies from follow students . British Journal of Educational Psychology , 59, 75–85. <http://dx.doi.org/10.1111/j.2044-8279.1989.tb03078.x>

Weiner, B. (1980) A cognitive (attribution)–emotion-action model of motivated behavior: An analysis of judgments of help giving . Journal of Personality and Social Psychology , 39, 186–200. <http://dx.doi.org/10.1037/0022-3514.39.2.186>

Weiner, B. (1985) An attribution theory of achievement motivation and emotion . Psychological Review , 92, 548–573. <http://dx.doi.org/10.1037/0033-295X.92.4.548>

Weiner, B. (1986) An Attributional Theory of Motivation and Emotion . New York: Springer-Verlag. <http://dx.doi.org/10.1007/978-1-4612-4948-1>

Weiner, B. (1992) Human Motivation: Metaphors, Theory, and Research . Newbury Park, CA: Sage.

Weiner, B. (1995) Judgments of Responsibility: A Foundation for a Theory of Social Conduct . New York: Guilford.

Weiner, B. (2006) Social Motivation, Justice, and the Moral Emotions . Mahway, NJ: Erlbaum Press.

Weiner, B. (2007) Examining emotional diversity in the classroom: An attribution theorist considers the moral emotions . In P.A. Schutz, ed. and R. Pekrun (eds), Emotion in Education , pp. 75–88. San Diego: Academic Press. <http://dx.doi.org/10.1016/B978-012372545-5/50006-X>

Weiner, B., Frieze, I.H., Kukla, A., Reed, L., Rest, S. and Rosenbaum, R.M. (1971) Perceiving the Causes of Success and Failure . Morristown, NJ: General Learning Press.

Weiner, B., Graham, S. and Chandler, C.C. (1982) Pity, anger, and guilt: An attributional analysis . Personality and Social Psychology Bulletin , 8, 226–232. <http://dx.doi.org/10.1177/0146167282082007>

Weiner, B. and Kukla, A. (1970) An attributional analysis of achievement motivation . Journal of Personality and Social Psychology , 15, 1–20. <http://dx.doi.org/10.1037/h0029211>

Weiner, B., Nierenberg, R. and Goldstein, M. (1976) Social learning (locus of control) versus attributional (causal stability) interpretations of expectancy of success . Journal of Personality , 44, 52–68. <http://dx.doi.org/10.1111/j.1467-6494.1976.tb00583.x>

Weiner, B., Perry, R.P. and Magnusson, J. (1988) An attributional analysis of reactions to stigmas . Journal of Personality and Social Psychology , 55, 738–748. <http://dx.doi.org/10.1037/0022-3514.55.5.738>

Weiner, B. and Potepan, P.A. (1970) Personality correlates and affective reactions toward exams of succeeding and failing college students . Journal of Educational Psychology , 61, 144–151. <http://dx.doi.org/10.1037/h0028915>

Weiner, B., Russell, D. and Lerman, D. (1978) Affective consequences of causal ascriptions . In J.H. Harvey, ed. , W.J. Ickes, ed. and R.F. Kidd (eds), New Directions in Attribution Research , 2, 59–88. Hillsdale, NJ: Lawrence Erlbaum Press.

Weiner, B., Russell, D. and Lerman, D. (1979) The cognition-emotion process in achievement-related contexts . Journal of Personality and Social Psychology , 37, 1211–1220. <http://dx.doi.org/10.1037/0022-3514.37.7.1211>

Wilson, T.D., Damiani, M. and Shelton, N. (2002) Improving the academic performance of college students with brief attributional interventions . In J. Aronson (ed.) Improving Academic Achievement: Impact of Psychological Factors on Education , pp. 89–108. San Diego: Academic Press. <http://dx.doi.org/10.1016/B978-012064455-1/50008-7>

<http://dx.doi.org/10.4135/9781446249215.n8>

Handbook of Theories of Social Psychology: Volume 1

A Theory of Social Information Processing

Contributors: Robert S. Wyer

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "A Theory of Social Information Processing"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n9>

Print pages: 156-178

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n9>

[p. 156 ↓]

Chapter 8: A Theory of Social Information Processing

Robert S. Wyer, Jr., ed.

Abstract

A general theory of social information processing was developed 25 years ago in an attempt to integrate diverse phenomena that I and others had identified in our research on social cognition. After several iterations, the model continues to provide a useful theoretical framework for conceptualizing and integrating the deliberate and automatic processes that occur at different stages of cognitive functioning, including comprehension, memory storage and retrieval, inference, judgment, decision making, and output generation. More specific formulations of belief organization and change, person impression formation, dual processing, and the interplay of goal-directed and unconscious processing can be viewed in terms of the cognitive operations that are governed by the general model I proposed.

Introduction

My research during much of the past 25 years has been guided by a general theoretical formulation of social information processing. The model specifies several stages of processing en route to a judgment or behavioral decision (comprehension, organization in memory, inference, integration, and the generation of an overt response). The processing at these stages can be either deliberative or automatic. More specific theoretical formulations of comprehension, inference, judgment, and decision making

can often be conceptualized in terms of the cognitive operations that occur at different stages of processing and the factors that influence their activation and use.

The theory has gone through several iterations (Wyer and Srull, 1980, 1986, 1989) en route to its present version (Wyer, 2004). In this chapter, I summarize the essential features of the model and discuss its potential for integrating the implications of other, more circumscribed formulations of information processing. The question is what motivated me to develop such a grandiose and all-encompassing model. The answer has its roots in my graduate work nearly 50 years ago.

[p. 157 ↓]

A Personal History

Initial Influences

Having been discouraged by my father from careers as a sports writer and jazz musician and not knowing what else to do, I wound up going to an engineering school. However, I felt intellectually stifled by my first job at Bell Laboratories and wanted to get a liberal arts education. I saw psychology as an indirect means of attaining this objective. When I applied to graduate school in psychology at the University of Colorado, I knew nothing at all about the field.

Research at Colorado during the early 1960s, largely stimulated by O.J. Harvey and Bill Scott, was focused on global structural characteristics of the cognitive system that might influence responses to social experience. During a summer job at Hughes Aircraft Company, I developed a model of artificial intelligence that provided measures of these constructs, and in my dissertation I attempted to relate the measures to various indices of personality and behavior. The research was actually published. However, my first job was at the University of Iowa. The Hull-Spence learning tradition that pervaded the atmosphere at Iowa during that period, coupled with some blistering reviews of some articles I had submitted for publication, convinced me that the approach I had been taking to understanding cognitive structure and function was a dead end.

The predominant view at Iowa during those early years was that thinking was an epiphenomenon of little relevance to a scientific investigation of human behavior. This view temporarily dampened my enthusiasm for studying cognitive processes. After leaving Iowa for the University of Illinois at Chicago Circle, however, four quite unrelated experiences rekindled my interest. First, I read a paper by Fishbein and Hunter (1964) on additivity and averaging processes in impression formation. Although I was unimpressed by this particular paper, I believed that the authors' theory was correct and conducted an even less impressive, never-to-be-published study to establish this fact. However, I sent the results of the study to Norman Anderson for comments and, to my surprise, he was quite encouraging. Stimulated by this reinforcement, I submitted a grant proposal to the National Science Foundation that was probably reviewed by Anderson and, therefore, was approved for funding. This resulted in a number of studies in the Anderson tradition which attempted to understand how people integrated the implications of personality traits into an overall person impression.

Second, a graduate course I was teaching stimulated me to read about W.J. McGuire's (1960) syllogistic model of belief organization. I realized that a slight modification of this model was consistent with the more general assumption that beliefs might be organized in memory in a manner that conformed to the laws of mathematical probability. Furthermore, if this were so, quantitative predictions could be made of the effects of change in one belief on other, unmentioned beliefs that were theoretically related to it. A flurry of experiments (for summaries, see Wyer and Hartwick, 1980; Wyer and Srull, 1989) evaluated implications of this possibility and of McGuire's formulation more generally.

Third, I happened to have an office next to Harry Upshaw. I became enamored of his conceptualization of how the perspective that people brought to bear on the stimuli they were judging affected their subjective positioning of the response scale they used to report their judgments (Upshaw, 1965). This conceptualization, which had implications for the effects of one's own attitude on responses to attitude-related messages, led to research in this area as well (Wyer, 1969).

Finally, I was asked to review the *Handbook of Personality Theory and Research*, which included another awe-inspiring chapter by McGuire (1968). In this chapter, McGuire conceptualized the different stages of processing that underlie responses to

a persuasive message and how the processing at each [p. 158 ↓] stage combined to influence the message's impact. This chapter led me to reflect upon my own work and to realize that, in fact, I had done work bearing on four different phases of processing: the organization of cognitions in memory, the use of previously formed cognitions to make an inference, information integration processes, and output (response) processes. I had simply not thought of them together. The outcome was my first book (Wyer, 1974). This book argued that human information processors were analogous to electronic processors (computers). Thus, they could simply be understood in terms of the "programs" they used to attain objectives specified by the user (others to whom they communicated in the social environment). My approach was rather naïve. However, it set the stage for my general interest in several different phases of information processing and how they might fit together.¹

The Advent of Social Cognition

Although the aforementioned experiences contained the seeds of my general interest in social information processing, it was not until the mid-1970s that this interest began to bear fruit. The conditions that stimulated the information processing model that I ultimately developed were quite unexpected and somewhat ironic.

In 1974, Norman Anderson and Seymour Rosenberg held a three-week workshop at the University of California, San Diego. Its purpose was to bring together research and theory on cognitive structure (exemplified by multidimensional scaling models) and process (represented by Anderson's formulations of functional measurement and information integration; see Anderson, 1971). The workshop was attended by several researchers, notably Tom Ostrom, Dave Hamilton and myself, who had been conducting impression formation research within the framework that Anderson proposed, but were frustrated by our failure to come to grips with the processes we intuitively believed to be involved in the formation of person impressions outside the laboratory. Reid Hastie and Ebbe Ebbesen, who had quite different perspectives on impression formation processes, were also participants.

Although we had all been reviewing one another's work for journals, few of us had ever met. We nevertheless found ourselves involved in daily discussions, often extending into the early hours of the morning, which culminated in our agreement that the research we had been doing was going nowhere. To make progress, we would need to understand how complex bodies of information were organized and stored in memory and how this information was later retrieved and operated upon in order to make a judgment.

To this end, the five of us, eventually joined by Don Carlston, began meeting informally twice a year over a long weekend, exchanging research findings and their implications.² Several findings, although now widely accepted, were nonintuitive at the time. For example, people were better able to recall someone's behaviors if they had received them for the purpose of forming an impression of the person than if they had been told explicitly to learn and remember the behaviors (Hamilton et al., 1980). People who had formed a trait-based impression of a person had better recall of the person's behaviors that were inconsistent with this impression than behaviors that confirmed its validity (Hastie and Kumar, 1979). This was not the case, however, when people formed impressions on the basis of observed behavior (Cohen and Ebbesen, 1979). Finally, once people had made one judgment of a person on the basis of trait descriptions, subsequent judgments of him were based on their first judgments regardless of the implications of the original information they had received (Lingle and Ostrom, 1979). Furthermore, the effects of initial judgments on later ones increased with the time interval between the two judgments (Carlston, 1980; Srull and [p. 159 ↓] Wyer, 1980). These and other findings were sufficiently provocative that we decided to publish them in a book on person memory (Hastie et al., 1980).

The Birth of a Theory

I was assigned to write an integrative chapter for the Hastie et al. volume that would bring others' diverging findings together within a single conceptual framework. I embarked on this project with Thorn Srull, a graduate student at the time. The challenge was rather daunting, as it required an integration of not only phenomena such as those mentioned earlier but also findings that had begun to emerge on the impact of cognitive

heuristics (e.g., Nisbett and Ross, 1980), knowledge accessibility (Higgins et al., 1977), and the earlier work we had done within the frameworks proposed by Anderson and McGuire. To account for Hamilton et al.'s (1980) findings, for example, we needed to take into account the effects of different processing goals on the way that information was organized in memory and was subsequently retrieved. To account for the effects of making an initial judgment on later ones (Carlston, 1980), we needed to develop a conceptualization of memory and judgment that would specify when and why one type of information took priority over another. We had to account for more specific encoding and organizational processes of the sort identified by Hastie and Kumar (1979). Finally, we needed to take into account the process of encoding ongoing behavior (Cohen and Ebbesen, 1979) as well as verbal information and the use of these nonverbal encodings on memory and judgment.

Our main challenge was to specify a common set of memory storage and retrieval processes that would apply at different stages of processing (comprehension, integration, inference, etc.) and would account for phenomena that were localized at these stages. Although Don Carlston and I had proposed an associative network model of the representation of social information in memory (Wyer and Carlston, 1979), it failed to capture the sort of findings that Hamilton, Hastie, Ostrom, and others were identifying. At the end of one of our early discussions, Srull incidentally made an observation that proved to be the central feature of our conceptualization: "If anyone is going to develop a viable conceptualization of social memory," he remarked, "they will have to drop the associative network metaphor." Having struggled with the Wyer and Carlston model, I quickly agreed. The result was the development of a "storage bin" metaphor of memory. That is, we conceptualized long-term memory as consisting of a number of content-addressable storage bins, into which different representations of knowledge could be deposited in the order they were formed and used, and from which representations could be retrieved on the basis of a probabilistic top-down search. The "bin" metaphor proved to be a powerful conceptual tool.

With this construct, we developed a metaphorical flow diagram consisting of a number of processing units (each pertaining to a different stage of processing) and memory storage units, all of which were interconnected by pathways that indicated the transmission of information from one unit to another. This was more challenging than we had expected. Any given pathway between two units had implications for

the type of information that could be transmitted and the conditions in which it could occur. At the same time, the failure to specify a path placed constraints on system functioning. Several iterations were required to come up with a model that could account for the phenomena we needed to explain without creating problems elsewhere in the processing system. As a consequence, what was originally intended solely as a descriptive device for integrating the phenomena reported elsewhere turned out to be a full-blown theory of social information processing with testable empirical implications.

[p. 160 ↓]

Further Iterations

The 1980 model was updated several times to account for new phenomena and to eliminate ambiguities that existed in its initial formulation (Wyer and Srull, 1986, 1989). After proposing the 1989 version of the model, however, I had a choice between continuing to test specific implications of the formulation and exploring new empirical issues that I found to be intrinsically interesting without considering their relevance to the model we had developed. As Bob Zajonc once observed, theories have many attackers, and unless someone is willing to defend them, they are likely to die a premature death. Often, however, the only person who is motivated to defend a theory is the individual who proposed it.

Nevertheless, I decided to let the model defend itself and to explore new horizons. This led me into research and theorizing on the antecedents of humor elicitation (Wyer and Collins, 1992), the influence of affect in information processing (Wyer et al., 1999), the representation of event sequences (Wyer et al., 2002), and comprehension processes more generally (Wyer and Gruenfeld, 1995; Wyer and Radvansky, 1999). In pursuing this latter line of research, however, it became clear that the original Wyer and Srull (1989) model could not account for the comprehension of information about familiar persons and events without substantially modifying its assumptions. This created a need (in my own mind, if not others') to modify the conceptualization in ways that would provide this account. The "final" (or, at least, most recent) version of the model, reported in Wyer (2004; Wyer and Radvansky, 1999) was the result.³

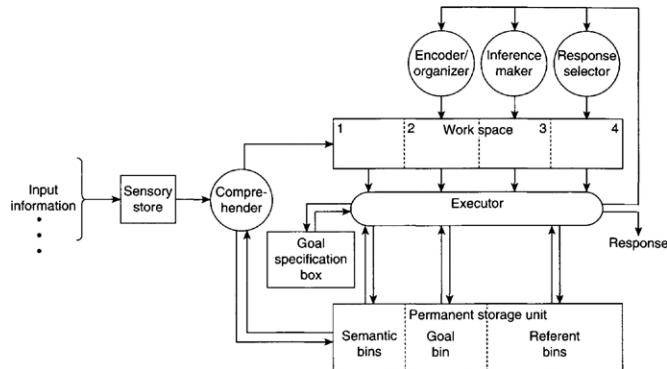
The Theory

The conceptualization of information processing that we developed is multifaceted, and a detailed exposition of it is beyond the scope of this chapter. Although it is capable of generating many specific predictions (e.g., Wyer and Srull, 1989), its primary value lies in its ability to provide a framework for integrating the phenomena identified in other, more specific areas of social information processing research. We hoped to account for the activities that occur at different stages of cognitive processing (comprehension, organization of information in memory, inference and integration, and the generation of overt responses) and to specify the processes of storing and retrieving information required at each of these stages. We further needed to address (1) the interplay of deliberative and automatic processing, (2) the unconscious influences of information on judgments, (3) the effects of externally imposed and internally generated goals on information processing at different stages, (4) the cognitive activities that occur in the absence of any conscious goal-directed activity (e.g., the free flow of thought), and (5) the theoretical relation between memory and judgment.

System Architecture

As shown schematically in [Figure 8.1](#), the model is composed of three major memory units, four special purchase processing units, and an Executor that directs the flow of information between these units. The memory units include the *Work Space* (analogous to working memory), the *Permanent Storage Unit* (long term memory), and a *Goal Specification Box*. The latter unit is a temporary repository of goal schemas, or sequences of cognitive steps that are involved in the pursuit of a particular objective.

Figure 8.1 Metaphorical representation of the Wyer and Srull (1989) model. Rectangles denote storage units. Ovals and circles denote processing units. Arrows denote the direction of transmission of cognitive material between these units



The processing units include a *Comprehender* which is an initial comprehension device that interprets input information in terms of general verb, noun, and adjective concepts, an *Encoder/Organizer*, which performs high-order interpretations of information and forms mental representations consisting of several features, an *Inference Maker*, which combines the implications of information to form a subjective [p. 161 ↓] judgment, and a *Response Selector*, which transforms subjective inferences into an overt response. The Comprehender comes into play in the processing of all of the information one receives and performs its functions automatically. The other units come into play only in the course of more specific, goal-directed processing. The activation of these latter units is governed by an *Executor*, which directs the flow of information between processing units and storage units. In doing so, it takes instructions from a goal schema that specifies the sequence of operations to be performed and the type of information required.

Systems Operation

The general operation of the processing system can be described briefly.

- 1. Input information is initially encoded by the Comprehender in terms of semantic concepts that are retrieved from Permanent Storage. The results of this processing are then transmitted to the Work Space.
- 2. If the new information contains a goal specification, the Executor retrieves a relevant goal schema from Permanent Storage and deposits it in the Goal

Specification Box. Based on the steps specified in the schema, it retrieves features of the input information along with previously acquired concepts and knowledge retrieved from Permanent Storage, and transmits them to a relevant special-purpose processing unit with an instruction to negotiate the step in question. The results of this processing are then returned to the Work Space.

- 3. The Executor again consults the relevant goal schema and, if no further processing is required, transmits the output of the initial processing to an additional processing unit along with the necessary goal-relevant knowledge required to perform its function. This continues until either the sequence of steps required to attain the objective is completed or other processing demands interfere.
- 4. *Simultaneous goal-directed processing.* More than one goal schema can be contained in the Goal Specification Box at any one time. This means that more than one goal can be pursued simultaneously. However, Its capacity is limited. Consequently, when the schema that is required to attain a particular objective is complex and detailed, there is less room for other schemas and so fewer other goals can be simultaneously pursued.
- 5. *Free flow of thought.* If no goal schema is activated, the system enters into a default routine. Specifically, the Executor samples a subset of features that are present in the Work Space and uses them as retrieval cues to identify a previously acquired knowledge representation. It then repeats this process, sampling a second set of features (which may include those of the representation just retrieved), and continuing until a goal specification is identified (based either on external input information or material retrieved from memory), in which case steps 1 and 2 are performed. Thus, this activity accounts for the free flow of thought that occurs in the absence of any specific goal-directed activity.

One other aspect of the model is important. That is, *consciousness resides in the Executor*. The Executor has access to (1) the information transmitted by the Comprehender, (2) previously acquired knowledge that is retrieved from Permanent Storage, (3) the steps involved in attaining a particular goal that are contained in a goal schema, and (4) the output of special-purpose processing units. Consequently, this material is subject to conscious awareness. However, the activities of the

Comprehender and other processing units, which are governed by routines that are stored in the libraries of these units, are *not* subject to awareness. Furthermore, note that the Comprehender, unlike other processing units, is not under the control of the Executor (see [Figure 8.1](#)). Thus, the system is not only unaware of the cognitive processes that underlie the comprehension of new information, but the processes are themselves not under conscious control. Thus, for example, it is impossible *not* to comprehend the statement “the boy kicked the ball” in terms of verb and noun concepts that are retrieved from permanent memory.⁴

Similarly, the actions performed by special-purpose processing units are not subject to awareness. Only the more general sequences of cognitive steps contained in a goal schema are conscious. Thus, suppose a goal is to form an impression of someone on the basis of information about the person's behaviors. The goal schema retrieved for use in attaining this objective might contain instructions to interpret the behaviors in terms of more general trait concepts and to form an evaluative concept of the individual on the basis of the behaviors' evaluative implications. These trait encodings and the overall evaluation, which are outputs of these processes, are subject to awareness. However, the specific cognitive mechanisms that are necessary to accomplish these steps are specified in the libraries of the processing units involved and are not subject to awareness.

Storage and Retrieval Processes

Three memory units are postulated. One, the Goal Specification Box, has already been discussed. The Work Space and Permanent Storage Unit require some elaboration.

Work Space

This unit is a temporary store of (1) information that has recently been transmitted to it by the Comprehender, (2) previously acquired concepts and knowledge that have been used by special-purpose processing units to attain specific processing objectives, and (3) the results of processing by these units. Information is retained in the Work Space

as long as it is likely to be involved in attaining goals that exist at the time. However, if an objective has been attained, or if information has not been employed for some time, the Work Space may be cleared to facilitate the identification of material that is of more immediate relevance. In this case, features of the original input information that have not been involved in higher order goal-directed processing (and, therefore, are not contained in a knowledge representation [p. 163 ↓] that is transmitted to Permanent Storage) are irretrievably lost.

Permanent Storage Unit

In the 1989 version of the model, permanent storage was assumed to consist of a number of content-addressable “storage bins.” One, *semantic* bin, contains general semantic concepts that are not specific to any particular referent and are drawn upon by the Comprehender in interpreting input information at an early stage of processing. A second, *goal* bin contains goal-schemas that are retrieved by the Executor for use in guiding goal-directed activity. The remaining, *referent* bins contain material pertaining to specific referents of goal-directed processing (persons, objects, events, or situations). Each referent bin is assigned a header, or set of verbal and nonverbal features that identify its referent. These features can include the name of the referent, a visual image, and attributes that have become strongly associated with the referent through learning and, can serve to identify it. A bin's referent can be at different levels of generality (e.g., George W. Bush, incompetent US Presidents, US Presidents, etc.). The contents of a bin can also vary. Thus, it can include single trait concepts that have been assigned to the referent, a verbally or nonverbally coded sequence of events that could describe either a prototypic experience (e.g., eating at a restaurant) or a specific one, a cluster of traits or interrelated behaviors, and a prior judgment or evaluation. Each such representation constitutes a separate unit of knowledge that is stored independently of other representations.

Five principles govern the storage and retrieval of these representations.

- 1. Units of knowledge are stored in a bin that refers to the stimulus to which processing objectives pertain. They are stored in the order they are formed, with the most recently deposited unit on top.

- 2. When information relevant to a referent is sought, the Work Space is the first location searched.
- 3. If no relevant information is found in the Work Space, a referent bin is identified on the basis of a set of probe cues that specify the nature of the referent and its header is reviewed for features that may be sufficient to attain the objective at hand. If these features are found, a search of the bin is not performed. (This means that if a feature has become strongly associated with a referent, it may be used as a basis for judgment independently of any more specific information whose implications might contradict it.)
- 4. If the features of a bin header are not sufficient to attain the objective at hand, a probabilistic top-down search is performed for a representation that is potentially relevant, and a copy of this representation is retrieved. Once it has been used, the copy is returned to the top of the bin. The probability of identifying a given unit of information in the top-down search of the bin is less than one. This means that the likelihood of retrieving a particular representation in the course of searching a bin is a function of both the recency of its acquisition and use (which determines its proximity to the top of the bin) and the frequency of its use (which determines the number of copies of it that have been formed and deposited there).
- 5. As noted earlier, information is deposited in a bin that is relevant to the processing objectives that exist at the time. Thus, if information is received that John kissed Mary and a goal is to form an impression of Mary, the information would be stored in a bin pertaining to her but not to a bin pertaining to John. This means that the information is unlikely to be retrieved later if information about John is required.

These retrieval processes are theoretically invoked at all stages of processing. Thus, they govern the concepts that are retrieved for use in both interpreting new information at the time it is received and later, in the pursuit of more specific objectives (e.g., to interpret a behavior in terms of a trait concept, to infer the likelihood of an event, to report an attitude toward a person or object, etc.). Furthermore, when alternative procedures can be used to attain a certain objective, these same processes govern the goal schema that is retrieved from the goal bin and applied.

The aforementioned principles have numerous implications (Wyer and Srull, [p. 164 ↓] 1989), one of which may be worth mention. Note that all input information that is

interpreted by the Comprehender is transmitted to the Work Space regardless of its relevance to any specific processing objective that might exist. However, only features of this information that are involved in goal-directed processing are transferred to Permanent Storage. Because the Work Space is not cleared immediately, features of the input information that are irrelevant to the goal for which the information was initially used can potentially be identified and brought to bear on other judgments and decisions that are made a short time after the information has been received. After a period of time has elapsed, however, only the results of the original goal-directed processing, which have been transmitted to Permanent Storage, can be used. This means that judgments and decisions that are made at the time information is first received can influence later judgments independently of the implications of this information. Furthermore, this influence increases over time.

Several studies support these predictions (e.g., Carlston, 1980; Lingle and Ostrom, 1979; Srull and Wyer, 1980, 1983). Carlston (1980), for example, found that people who had judged a person's honesty on the basis of a behavior that were both honest and unkind (e.g., turning in a friend for cheating on an exam) later judged him as relatively kind, whereas people who had initially judged his kindness later judged him as relatively dishonest. Moreover, this difference was greater three days after the first judgment was made than it was immediately afterwards. Research by Srull and Wyer (1980) has similar implications.

The Role of Procedural Knowledge

The processes described in the preceding section govern the storage and retrieval of *declarative* knowledge that individuals acquire. This knowledge presumably includes the goal schemas that are formed for the purpose of attaining a particular objective. These schemas, which presumably consist of a sequence of events similar to those that compose a script, function as mental "recipes" that people intentionally consult for information about the steps required in order to attain a goal.

However, the theory distinguishes between these knowledge representations and the routines that compose the libraries of processing units. These routines, which constitute *procedural* knowledge, are acquired through learning and, once learned,

are theoretically applied without conscious deliberation. These routines can be conceptualized as “If [X], then [Y]” *productions* similar to those conceptualized by J. Anderson (1983; see also Smith, 1990), where [X] consists of a configuration of internally generated or externally impinging stimulus features and [Y] is a sequence of cognitive or motor behaviors that is activated and applied automatically when the eliciting conditions [X] are met. Thus, the configuration of features that compose [X] can include both representations of external stimuli that impinge on the cognitive system and internally generated stimuli (thoughts, proprioceptive reactions, feelings, etc.). However, the configuration is responded to as a whole. Therefore, not all of its individual features may be consciously identified in order for the behavior routines associated with them to be activated. Thus, although these routines may be activated by components of declarative knowledge, they may be performed without recourse to this knowledge.

Comprehension Revisited

Although the 1989 version of the model was successful in accounting for a variety of phenomena (Wyer and Srull, 1989), subsequent research we performed made salient a number of deficiencies, most of which were traceable to problems at the early, [p. 165 ↓] comprehension stage of processing. First, the Comprehender had access to only the semantic bin in Permanent Storage (see [Figure 8.1](#)). This implies that new stimulus information is first interpreted in terms of general semantic concepts and that referent-specific concepts are not applied until a later stage of processing. However, Wyer and Radvansky (1999) found that statements about well-known individuals and their behavior (e.g., “Jane Fonda did aerobics”) were comprehended more quickly than equivalent statements about members of the semantic categories to which the individuals belonged (e.g., “The actress did aerobics”). Furthermore, the time required to comprehend statements about known persons was not appreciably different from the time required to verify them as true or false.

Second, the 1989 model failed to take into account the role of visual imagery. Radvansky et al. (1997) provided evidence that individuals who encounter statements about objects or events that are situationally and temporally constrained construct a mental simulation, or *situation model*, of the events that consists in part of a

mental image whose features are spatially and temporally organized. Although the Comprehender was assumed to be capable of processing information in different sense modalities, the unique contribution of visual imagery in comprehending verbal information was not considered.

To account for these phenomena, the 2004 version of the model relaxed the assumption that the Comprehender had access to the semantic bin alone and allowed it to access referent-specific knowledge as well. This change not only permitted new information to be comprehended spontaneously in terms of specific concepts of its referents but also allowed for visual images to be constructed. It also allowed for the possibility that one's recognition of the validity of referent-specific information is often an inherent component of comprehension that occurs automatically rather than as a result of deliberative goal-directed cognitive activity.

Briefly, we assumed that a verbal description of a referent's observable behavior is comprehended by retrieving previously formed visual images of both the behavior and the actor and combining these images to form a new situation model that includes both. If this model is sufficiently similar to a preexisting situation model that is identified in the course of this processing, the information is spontaneously recognized as true, and if it is below a minimal threshold of similarity, it is spontaneously recognized as false. (Otherwise, the information is simply comprehended and the model of it is stored without construing its validity.)

Two implications of these processes are noteworthy. First, behaviors of a well-known person are likely to be comprehended with reference to a previously formed situation model that exemplifies them. In contrast, behaviors of an unknown person are more likely to be interpreted with reference to a more general, prototypic event representation. Colcombe and Wyer (2002) confirmed this difference.

Second, although information that is conveyed verbally may be spontaneously coded visually in the course of comprehending it (Jiang and Wyer, 2009), information that is transmitted visually is comprehended with reference to previously formed situation models but may *not* be spontaneously encoded verbally. This latter recoding may occur only if a verbal coding is necessary to attain some more specific processing objective (Adaval and Wyer, 2004; Wyer et al., 2002).

The spontaneous verification processes assumed by the 2004 extension of the model also permit the model to account for other important comprehension phenomena, such as the spontaneous reactions to statements that violate normative principles of communication (e.g., the principles that communications are typically expected to be informative and to convey the truth as the communicator sees it; Grice, 1975). That is, violations of these rules are spontaneously [p. 166 ↓] recognized and stimulate higher order comprehension processes in order to construe the communication's intended meaning.

Costs and Benefits

As the preceding examples indicate, the 2004 version of the model can account for several phenomena that could not be explained by the 1989 version. However, the revision has had costs as well as benefits. The costs have come as a result of the need to revise the model's assumptions about memory storage and retrieval. That is, the original model assumed that information about specific referents is retrieved from Permanent Storage by first identifying a referent bin on the basis of features contained in its header and then performing a sequential, top-down search of the bin's contents for goal-relevant information. However, the time required for this two-stage process would contradict the evidence that referent-specific information is comprehended very quickly in the absence of specific goal-directed processing. Consequently, the 2004 version of the model required a relaxation of the "bin" construct and the postulation of a memory retrieval model similar to that proposed by Ratcliff (1978), which makes few assumptions about the organization of concepts and knowledge in memory. I am reluctant to discard the bin metaphor completely, as it is a powerful tool in conceptualizing goal-directed cognitive activity at later, postcomprehension stages of processing. However, the two sets of retrieval assumptions have not been fully reconciled at this writing.

The Nature of Goal Schemas: Two Examples

The Wyer and Srull conceptualization is obviously an incomplete description of information processing in the absence of a more precise statement of the goal schemas that govern the attainment of specific processing objectives. However, more circumscribed theoretical formulations can often be conceptualized in terms of the content of these goal schemas. Much of my work in specific areas of social information processing can be viewed as attempts to specify the general nature of these schemas.

For example, the information integration processes that were the focus of my early research in impression formation can be viewed as hypotheses concerning the routines that are stored in the Inference Maker's library and used to combine the implications of individual pieces of information to form a judgment. As such, they are not subject to conscious awareness. On the other hand, the particular routine that is applied may depend on both the type of information presented and the situational context in which the judgment is made (Wyer and Carlston, 1979). These latter contingencies may be specified in the goal schema that governs the processing of the information and, therefore, in the instructions the Executor gives to the Inference Maker along with the information to be integrated. Two other formulations we developed, each of which can be viewed as a theory of cognitive functioning in its own right, provide more specific examples of goal schemas.

Belief Formation and Change

My extension of McGuire's (1960) syllogistic model of belief organization, mentioned earlier, assumes that if beliefs (estimates of the likelihood that a proposition is true) are defined in units of probability, the relation between the belief in a conclusion (C) and beliefs in an antecedent (A) can be described by the equation:

$$P(C) = P(A)P(C/A) + P(\sim A)P(C/\sim A) \quad (8.1)$$

where $P(C)$ is the belief that C is true, $P(A)$ and $P(\sim A)$ are beliefs that A is and is not true, [p. 167 ↓] respectively, and $P(C/A)$ and $P(C/\sim A)$ are beliefs that C is true if A is and is not true, respectively. Furthermore, if information is received that changes beliefs in A by an increment $\Delta P(A)$, its effect on the peripheral belief in C is simply,

$$\Delta P(C) = \Delta P(A)[P(C/A) - P(C/\sim A)] \quad (8.2)$$

These two equations provide remarkably accurate quantitative fits of the relations among the beliefs involved (Wyer, 1970). Furthermore, they capture the Socratic effect, or the tendency for related beliefs, once they are made salient in temporal proximity, to become more consistent over time (Rosen and Wyer, 1972). However, although this conceptualization was originally interpreted as a model of belief organization, subsequent research (Wyer and Hartwick, 1980) indicated that it was more appropriately conceptualized as a description of *inference* processes, namely, the process of inferring the likelihood of C from beliefs in a second, “informational” proposition, A , that happens to be salient at the time. For example, individuals who are asked their belief that drinking coffee is desirable (C) are likely to search memory for a second, informational proposition, A , that has implications for it and to construe the implications of this proposition. Thus, they may draw a different conclusion if the proposition they happen to retrieve is, “Drinking coffee wakes you up in the morning” than if it is, “Drinking coffee gives you insomnia” (Wyer and Hartwick, 1980). The proposition they happen to identify and use presumably depends on its accessibility in the referent bin in which it is located.

This process, which presumably is stored in the library of the Inference Maker (see [Figure 8.1](#)) is applied as a result of instructions by the Executor to make an inference about C based on information pertaining to A . To this extent, the process described by Equation (8.1) may be performed without awareness.

Impression Formation

A second example of a goal schema is provided by our conceptualization of person memory and judgments (Srull and Wyer, 1989; Wyer and Srull, 1989). The

conceptualization specifies the sequence of cognitive activities that occur when individuals have the goal of forming an impression of someone on the basis of a series of traits and behaviors, the judgments that are based on this representation, and the subsequent recall of the behaviors contained in it. The theory assumes that people with an impression formation objective attempt (a) to assign general personality traits to the person on the basis of the information about him, and (b) to arrive at an overall evaluation of the person as likeable or dislikeable. To this end, the following steps are performed:

- 1. *Trait encoding.* People encode the individual behavior of the person in terms of trait concepts that they exemplify. This activity establishes an association between the behavior and the trait concept it exemplifies. When several behaviors exemplify the same trait, a trait-behavior cluster is formed.
- 2. *Evaluative concept formation.* The initial information presented about the person is coded evaluatively and, if its implications are consistent, a general concept is formed of the individual as likeable or dislikeable. The person's behaviors are also encoded evaluatively and are thought about with reference to this concept, consequently becoming associated with it.
- 3. *Responses to inconsistency.* If a behavior is encountered that is evaluatively inconsistent with the general person concept that is formed in "2," people respond in two ways. First, they think about the behavior in relation to others they have received in an attempt to reconcile its occurrence. This activity leads associations to be formed between the inconsistent behavior and others. Second, they review behaviors that are consistent with their general concept of the person in an attempt to confirm its validity. This activity strengthens the association between these behaviors and the concept.
- 4. The representations that are formed (both trait-behavior clusters and the more general person representation) are then stored in memory in a bin pertaining to the person.

[p. 168 ↓]

Retrieval Processes

Suppose that people are later asked to recall the information they have received about the person. They will retrieve one of the representations from the bin in which it is located and search its contents, beginning with the central node and progressing down one of the pathways emanating from it to a behavior. After reporting this behavior, they progress along a pathway to another behavior and so on, returning to the central node whenever they encounter a dead end. Thus, if the representation they happen to retrieve is a trait-behavior cluster, they are likely to recall the behaviors in this cluster before behaviors pertaining to other traits (for evidence, see Hamilton et al., 1980). If they retrieve and use the person representation, the first behavior they recall should be consistent with the evaluative concept of the person, as consistent behaviors are more strongly associated with the concept. In general, however, there are more pathways leading into inconsistent behaviors than consistent ones, and so inconsistent behaviors are better recalled (Hastie and Kumar, 1979; Srull, 1981).

Judgment

When people are called upon to make a judgment, they retrieve a representation whose central concept is relevant to the judgment and use its implications as a basis for the judgment without reviewing the behaviors that are associated with it. These behaviors are only consulted if the central concept has no direct implications for the judgment to be made. Thus, for example, if people are asked to evaluate the person, they should make judgments that are implied by the evaluative person concept they have formed. Yet, behaviors that are inconsistent with this concept may often be better recalled.

Qualifications

The aforementioned conceptualization is able to account for numerous memory and judgment phenomena, including the impact of stereotypes on memory (Wyer and Martin, 1986), differences in memory for individuals and groups (Srull, 1981; Wyer

et al., 1984), and the impact of information that people are told to disregard (Wyer and Budesheim, 1987). However, its applicability may nonetheless be limited to the particular paradigm that was normally used to evaluate its validity. When a person's behaviors are conveyed in the course of an informal conversation, for example, listeners tend to form an impression of the *speakers* rather than of the individuals they are describing and organize their descriptions of the person's behaviors around these impressions (Wyer et al., 1990). Thus, they think more extensively about behavior descriptions that violate conversational norms to be polite and modest (Wyer et al., 1994).

Despite these limitations, our original conceptualization of person's memory and judgment processes provides an example of a goal schema that might, in principle, be constructed and incorporated into the more general model that Wyer and Srull proposed. Furthermore, the qualifications on the original model that were identified in later research can be viewed as reflecting different goal schemas that are specific to different types of information and situational contexts.

Relation to other Theories

As our examples of goal schemas indicate, a primary value of the conceptualization lies in the framework it provides for integrating the implications of more specific formulations of information processing. Although space limitations preclude a full discussion of this possibility, a few examples may illustrate its implications of other formulations and its utility in identifying further avenues for empirical investigation.

Dual Processing Models

Numerous conceptualizations have been proposed that fall under the heading of "dual processing models." For some reason, the [p. 169 ↓] Wyer and Srull model has rarely, if ever, been formally recognized as such a model (Chaiken and Trope, 1999). As the preceding discussion indicates, however, it is actually a *multi*-process model that allows for several alternative processes at each stage of cognitive activity (comprehension, inference, etc.), some of which are automatic and others are deliberative. Furthermore,

it provides a framework for conceptualizing the conditions in which different processes occur.

A model proposed by Strack and Deutsch (2004) is particularly interesting to consider in this context. Their model postulates two separate processing systems. One, a *reflective*, system comes into play in deliberative goal-directed processing and is governed by processes of which individuals are well aware. Thus, it generates judgments, decisions, and intentions that depend on the particular type of goal at hand. A second, *impulsive*, system operates automatically and is governed largely by associative processes. This system directs behavior through cognition-behavior associations that are acquired through learning. Once acquired, however, their activation is governed by general principles of knowledge accessibility (Förster and Liberman, 2007; Higgins, 1996; Wyer, 2008).

The automatic processes that compose the “impulsive” system defined by Strack and Deutsch are localized in several components of the Wyer and Srull model, including the Comprehender and the libraries of special-purpose processing units. These processes are also implied in the activities of the Executor in the course of free flow of thought. There is an important distinction between the two conceptualizations, however; that is, the reflective goal-directed processing assumed by Strack and Deutsch appears to be completely governed by goal schemas that exist as part of general knowledge. The impulsive system only operates as a default, when conscious goal-directed actions performed by the reflective system are not operating. In contrast, the Wyer and Srull model allows for automatic (unconscious) processes to occur in the pursuit of conscious goal-directed activity. Specifically, the processes that are stored in the library of the various processing units that are activated by the model are goal-directed, but nonetheless operate automatically without consciousness of the specific cognitive operations that are involved.

Attitude Formation and Change

As Schwarz and Bohner (2001) point out, the conditions in which individuals consciously change their attitude or opinion in response to new information may be limited. More generally, individuals who are called upon to evaluate a stimulus may

retrieve a sample of judgment-relevant information that happens to be easily accessible in memory and compute a judgment online, based on the implications of this sample. Thus, the effect of a persuasive communication may not reflect a conscious change in attitude. Rather, it would result from a difference in the information that enters into the online computation of the attitudes at different points in time.

This conceptualization is quite consistent with the Wyer and Srull model. This model assumes that an attitude toward an object is computed on the basis of a subset of the information stored in a referent bin that has most recently or frequently been used. This information can sometimes include a previously formed judgment that has been stored in the bin as a result of prior processing. To this extent, whether the attitude appears to be stable depends on whether this judgment is the *only* information retrieved or other information is retrieved as well. This depends on the recency with which other judgment-relevant information has been acquired.

Construing the Implications of a Persuasive Message

The dominant theories of communication and persuasion have been proposed by Petty [p. 170 ↓] and Cacioppo (1986) and Chaiken (1987). Petty and Cacioppo's model assumes that individuals base their judgments on different criteria, depending on their a priori assessment of the time and cognitive resources that are required to apply them. In contrast, Chaiken's (1987) model assumes that judgmental criteria are applied sequentially, with easily accessible and easy-to-apply criteria considered first and additional criteria considered only if the judgment is important and participants have little confidence in the first criteria they apply.

Of the two conceptualizations, Chaiken's is more congenial to the Wyer and Srull formulation. The sequential identification and use of alternative judgment criteria, and the threshold of confidence required to make a judgment, are presumably governed by a goal schema, and the relative accessibility of the criteria being applied would be determined by the likelihood of identifying them in a top-down search of a referent bin pertaining to the communication's referent.

Both Chaiken's conceptualization and Petty and Cacioppo's, however, leave open the question of precisely how an inference is computed. As Kruglanski et al. (1999) point out, this process could be similar regardless of whether the inference is based on the source of a message or the arguments contained in it. In fact, the syllogistic process assumed by Kruglanski et al.'s (1999) "unimodel" is very similar to the process implied by the belief inference model described earlier in this chapter as an example of a goal schema (see Wyer, 2006, for an elaboration).

Dual and Implicit Attitudes

Several conceptualizations postulate the existence of "dual attitudes" that exist simultaneously in memory and are called upon for use in judgments or behavioral decisions, depending on the circumstances. Greenwald and Banaji (1995), for example, distinguish between *explicit* attitudes that individuals consciously report and *implicit* attitudes that mediate favorable thoughts or behavior toward a stimulus without necessarily any awareness of their influence or, in fact, their existence.

The existence of different attitude-based representations in memory is obviously compatible with the Wyer and Srull model, and their use as bases for judgment is presumably determined by goal schemas that are retrieved for use under the conditions at hand. For example, Brunei et al. (2004) showed that European Americans' explicit attitudes toward advertisements were similar regardless of the ethnicity of the models shown in the ads, but their implicit attitudes were more favorable when the models were white. The procedures used to assess these attitudes differed considerably. Thus, in terms of the Wyer and Srull model, these data simply suggest that the goal schema that individuals activate for use in computing a judgment depends on the type of judgment that they are asked to make.

Unconscious Influences on Judgments and Decisions

Perhaps the most provocative body of research to emerge in recent years has shown that individuals' judgments and behavior are often governed by factors of which they are not consciously aware (cf. Bargh et al., 1996; Chartrand and Bargh, 1996; for a review, see Dijksterhuis and Bargh, 2001). These studies suggest that the perception of an event can elicit behavior automatically with little, if any, interpolated cognitive activity. A conceptualization of these effects within the framework proposed by Wyer and Srull, however, suggests the need to distinguish between (a) consciousness of the conditions that lead concepts to become accessible in memory and (b) consciousness of the connection between these conditions and the behavior that the concepts elicit.

[p. 171 ↓] Specifically, the processing of information by the Comprehender is not under control of the Executor and, therefore, its activities are not subject to conscious awareness. Concepts retrieved by the Comprehender can therefore become more accessible without conscious awareness of either the concepts themselves or the processes that give rise to their activation. As a consequence, the concepts are likely to be used to comprehend information that is received later (Bargh and Pietromonaco, 1982).

For concepts to come into play at later, postcomprehension stages of processing, however, they must somehow find their way into the Work Space where they can be selected as part of the set of features that the Executor compiles and either (a) uses to identify a goal schema or (b) transmits to a special-purpose processing unit, thereby influencing the particular routine that is used to attain the objective at hand. If the concepts are primed in the course of conscious goal-directed activity, they may enter the Work Space even if participants are unaware of the relation of the priming task to the behavior they perform subsequently. If concepts are primed subliminally and are not involved in postcomprehension goal-directed activity, however, this may not be the case.

The implications of this contingency are exemplified by Bargh et al.'s (1996) finding that when participants had been unobtrusively primed with concepts of the elderly, they walked more slowly to the elevator upon leaving the experiment. In this study, however, concepts were primed using a sentence-construction task that required their use. This goal-directed processing may have led the concepts to be present in the Work Space and to be fortuitously sampled by the Executor and transmitted to the Response Selector, where it influenced the selection of a goal schema that was used to generate an overt response. Thus, the effects identified by Bargh et al. might not have occurred if elderly-related concepts had been primed subliminally.

Goal-Directed Behavior and Motivation

Kruglanski and his colleagues (Kruglanski et al., 2002; Shah et al., 2003) postulate a hierarchy of goals at different levels of specificity such that the goal at one level serves as a means of attaining goals at other levels. In support of this conceptualization, Kruglanski et al. show that subliminally activating one goal can increase the accessibility of other goals that are both higher and lower in the hierarchy. Furthermore, once goal-relevant concepts are activated, they not only facilitate goal-directed behavior to which they are relevant but also interfere with goal-directed activity to which they are not relevant (Shah and Kruglanski, 2002).

Wyer and Srull (1989) also assume that goal schemas are represented in memory at different levels of generality, and that the activation of a goal at a particular level of specificity can activate the sequence of actions that are means to its attainment. If several goal schemas can be used to attain the same objective, their accessibility in memory depends on the frequency or recency with which they have been applied.

The evidence that *subliminally* activating one component of a goal sequence can increase the accessibility of other components (Kruglanski et al., 2002) is difficult for the Wyer and Srull (1989) conceptualization to explain. The conceptualization can nevertheless account for unconscious goal-directed processing. The routines in a processing unit's library are presumably performed in the course of goal-directed activity without awareness of the general objective to which they pertain. Moreover, although goal-related concepts can be among the configuration of features that elicits

a production of the sort that is contained in a processing unit's library, the configuration is responded to schematically and, therefore, might activate the production without consciousness of the specific features that compose it. Thus, goal-directed behavior can often occur without awareness [p. 172 ↓] of the goal to which it is relevant (Chartrand and Bargh, 1996).

Applications

Many aspects of the information-processing model have implications for information processing outside the laboratory. The model's implications for the spontaneous recognition of statements as true or false in the course of comprehending them (Wyer and Radvansky, 1999) and the processes that underlie humor elicitation (Wyer and Collins, 1992) are described in detail elsewhere, and space does not permit an elaboration in this chapter. My recent research has not provided rigorous tests of the conceptualization but, rather, has evaluated its more general implications for nonlaboratory phenomena. Two general areas of research may be worth noting briefly.

The Role of Visual Imagery in Responses to Persuasive Appeals

To reiterate, the 2004 version of the theory assumes that situation models are typically formed in the course of comprehending situation-specific events, a central component of which is a visual image. Several implications of this assumption have been investigated.

Effects of Donation Appeals

Appeals for donations often encourage recipients to form a mental representation (or, in terms of the 2004 theory, a "situation model") of themselves in the situation confronting the individuals in need of help. To be effective, however, the appeal must also induce the recipients to imagine themselves as a potential donor. However, it is

difficult if not impossible to form mental images of oneself in two roles simultaneously. Thus, suppose recipients of an appeal have the perspective of a potential donor at the time they receive a request for aid. In this case, an appeal that stimulates them to construct a situation model of themselves in the role of a victim may create cognitive conflict and decrease the appeal's effectiveness. Consistent with this conjecture, Iris Hung (Hung and Wyer, 2009) found that stimulating individuals to imagine a situation from the victims' perspective increased the appeal's effectiveness when a request for aid was not made until after the situation was described (as indicated by the amount of money that participants were willing to give). When participants had a disposition to imagine themselves as the donor at the outset, however, stimulating them to form a representation of the situation from the victim's perspective *decreased* the appeal's effectiveness.

Effects of Advertising

Other effects of visual imagery on the impact of persuasive appeals were identified by Jiang (2008; see Wyer et al., 2008). In a representative study, participants received information about a hotel describing features that could be imagined from either the same visual perspective (i.e., inside the hotel) or different perspectives (e.g., both inside and outside). In the first case, individuals could easily form a single image-based situation model of the hotel. In the second case, however, this was impossible to do. Consequently, individuals with a disposition to form visual images evaluated the hotel more favorably in the first case than in the second. When individuals were induced to process the information semantically without forming visual images, however, this difference was eliminated.

Additional evidence that difficulty in constructing situation models can have an adverse effect on the impact of advertisements was obtained by Hung and Wyer (2008). Participants received a print ad for a product consisting of (a) a problem that the product purportedly remedied (hair loss, stained clothing, etc.) and (b) the consequences of using the product. When one component was conveyed in a picture and the [p. 173 ↓] other was described verbally, recipients could construct a situation model of the sequence of events that was consistent with their general knowledge of the type of product being promoted. As a result, they evaluated the product favorably. When both

components were pictured, however, flexibility in constructing a visual image of the events described was eliminated, and so a situation model of the events was more difficult. Consequently, the product was evaluated less favorably.

The Role of Procedural Knowledge in Shopping Behavior

A central feature of the information processing model surrounds its ability to conceptualize the effects of a number of different processing objectives on judgments and decisions. These effects theoretically depend on how the procedures for attaining these objectives are represented in memory. Sequences of goal-directed behavior can be stored in memory either as part of semantic knowledge or as a production of the sort postulated by Anderson (1983) that is applied automatically when the preconditions for its activation are met. The impact of these different types of knowledge representations on consumer behavior has been demonstrated by Jing Xu (Xu and Wyer, 2007, 2008) and Hao Shen (Shen and Wyer, 2008), respectively.

Xu, for example, assumed that purchases are often governed by a three-stage sequence of activities involving (a) whether to make a purchase, (b) which alternative to purchase, and (c) how to implement the purchase. However, she further assumed that if situational factors activate the second stage of this sequence, consumers might perform it and proceed to the third, implemental stage without engaging in the first stage. Therefore, they might be disposed to buy one of the alternatives without considering whether they wanted to buy anything at all.

Finally, note that the process of deciding which product to buy is a special case of a more general process of making comparative judgments. To this extent, a “which-to-buy” mindset might be activated by other exemplars of this process that have nothing to do with purchase behavior. This was in fact the case; asking persons to compare the physical attributes of wild animals, or to indicate whether one country was similar to another, was sufficient to activate a “which-to-buy” mindset and, therefore, to increase their likelihood of purchasing products that were on sale after the experiment.⁵ Further evidence that behavior in one stimulus domain can induce a mindset that affects

behavior in other, quite different domains (specifically, variety seeking) was obtained by Shen and Wyer (2009).

Conclusion

The general information processing model described in this chapter is obviously metaphorical and, as such, must be evaluated in terms of its utility and not its validity. In my own work, I have not treated the model as a definitive theory of cognitive behavior but have viewed it as a heuristic device that calls attention to theoretical and empirical issues that might otherwise have escaped my attention. In this regard, I believe that a model should be able to defend itself and that the theorist should be its harshest critic. In this spirit, however, I want to thank the model personally for the stimulation it has provided me over the years. If the model has been able to stimulate a few others as well, it should feel doubly honored.

Acknowledgments

The writing of this paper and some of the research reported was supported by grant GRF641308 from the Research Grants Council of the Hong Kong SAR, China.

[p. 174 ↓]

Notes

1 As implied earlier, this interest reflects my attempt to emulate W.J. McGuire, social psychology's intellectual and spiritual leader and the father of social information processing theory and research. Both his own work and the personal support and encouragement he gave to me throughout my career were an inspiration, and any success I may have had results in large part from my conscious and unconscious attempts to follow, albeit feebly, in his footsteps.

2 These meetings gradually expanded into what is now known as the “Person Memory Interest Group,” initially inspired by Tom Ostrom and continued by Dave Hamilton and Eliot Smith. It now has over 80 members and is the primary organization for researchers in social cognition.

3 I earlier acknowledged my indebtedness to W.J. McGuire, whose research and theorizing obviously provided the inspiration for my own work (see footnote 1). However, his most important influence is less obvious. In 1968, almost six years after completing my PhD, I was still floundering, and wondering whether anyone had even read any of the work I had done, to say nothing of whether they cared about it. During this period of self-doubt, I submitted a paper to the *Journal of Personality and Social Psychology*, which McGuire processed as editor. Although I cannot recall his specific reactions, he undoubtedly called attention to numerous instances conceptual and expositional sloppiness, because I wrote back apologizing to him for putting him through the ordeal of evaluating it. I immediately received a response that made my day and, ultimately my career. I cannot recall his exact words. However, they were to the effect that he and others had “assumed” that I knew I was a good psychologist and that it was only in this context that they bothered to take the time to “carp” about the things he had noted. Leaving aside the fact that his comments could hardly be viewed as “carps,” this was the first time that anyone, let alone someone as eminent as W.J. McGuire, had conveyed any interest whatsoever in anything I had done. His encouragement at this critical point in my career gave me the confidence to persist. Many years later, it still inspires me during times of disappointment and self-doubt. In the context of this personal history, I would like to acknowledge my indebtedness to this remarkable psychologist and equally remarkable human being.

4 On the other hand, the statement “the ball kicked the boy” cannot be automatically comprehended on the basis of routines in the Comprehender’s library. Such a statement would therefore be transmitted to the Work Space along with a message signifying its status and would be comprehended (if at all) on the basis of Executor-controlled processing.

5 An interesting speculation derived from this conceptualization is that the consumption of material goods is greater during election years, when citizens are continually being asked which of two political candidates they prefer, than in off-election years. In fact,

an analysis of years between 1953 and 2000 showed that total retail-store sales in the United States were 9.4 percent higher during the three months prior to an election (August, September, and October) than during comparable months in the years before and after election years. Although this difference was not statistically significant, its consistency with our laboratory findings is provocative.

References

- Adaval, R. and Wyer, R.S. (2004). Communicating about a social interaction: Effects on memory for protagonists' statements and nonverbal behaviors . *Journal of Experimental Social Psychology* , 40, 450–465. <http://dx.doi.org/10.1016/j.jesp.2003.08.001>
- Anderson, J.R. (1983) *The Architecture of Cognition* . Cambridge, MA: Harvard University Press.
- Anderson, N.H. (1971) Integration theory and attitude change . *Psychological Review* , 78, 171–206. <http://dx.doi.org/10.1037/h0030834>
- Bargh, J.A., Chen, M. and Burrows, L. (1996) Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action . *Journal of Personality and Social Psychology* , 71, 230–244. <http://dx.doi.org/10.1037/0022-3514.71.2.230>
- Bargh, J.A. and Pietromonaco, P. (1982) Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation . *Journal of Personality and Social Psychology* , 43, 437–449. <http://dx.doi.org/10.1037/0022-3514.43.3.437>
- Brunei, F.F., Tietje, B.C. and Greenwald, A.G. (2004) Is the Implicit Association Test a valid and valuable measure of implicit consumer social cognition? *Journal of Consumer Psychology* , 14, 385–404. http://dx.doi.org/10.1207/s15327663jcp1404_8
- Carlston, D.E. (1980) Events, inferences and impression formation . In R. Hastie, ed. , T. Ostrom, ed. , E. Ebbesen, ed. , R. Wyer, ed. , D. Hamilton, ed. , and D. Carlston (eds), *Person Memory: The Cognitive Basis of Social Perception* , pp. 89–119. Hillsdale, NJ: Erlbaum.

Chaiken, S. (1987) The heuristic model of persuasion . In M.P. Zanna, ed. , J.M. Olson, ed. , and C.P. Herman (eds), Social Influence: The Ontario Symposium , 5, 3–39. Hillsdale, NJ: Erlbaum.

Chaiken, S. and Trope, Y. (1999) Dual-process Theories in Social Psychology . New York: Guilford Press.

Chartrand, T.L. and Bargh, J.A. (1996) Automatic activation of impression formation and memorization goals: Nonconscious goal priming reproduces effects of explicit task instructions . Journal of Personality and Social Psychology , 71, 464–478. <http://dx.doi.org/10.1037/0022-3514.71.3.464>

Cohen, C.E. and Ebbesen, E.B. (1979) Observational goals and schema activation: A theoretical framework for behavior perception . Journal of Experimental Social Psychology , 15, 305–339. <http://dx.doi.org/10.1016/0022-1031%2879%2990041-6>

Colcombe, S.J. and Wyer, R.S. (2002) The role of prototypes in the mental representation of temporally-related events . Cognitive Psychology , 44, 67–103. <http://dx.doi.org/10.1006/cogp.2001.0766>

Dijksterhuis, A. and Bargh, J.A. (2001) The perception-behavior expressway: Automatic effects of social perception on social behavior . In M.P. Zanna (ed.), Advances in Experimental Social Psychology , 33, 1–40. San Diego: Academic Press.

Fishbein, M. and Hunter, R. (1964) Summation versus balance in attitude organization and change . Journal of Abnormal and Social Psychology , 69, 505–510. <http://dx.doi.org/10.1037/h0048533>

Förster, J. and Liberman, N. (2007) Knowledge activation . In A. Kruglanski, ed. and E.T. Higgins (eds), Social Psychology: Handbook of Basic Principles (2nd edition), pp. 201–231. New York: Guilford.

Greenwald, A.G. and Banaji, M.R. (1995) Implicit social cognition: Attitudes, self-esteem, and stereotypes . Psychological Review , 102, 4–27. <http://dx.doi.org/10.1037/0033-295X.102.1.4>

Grice, H.P. (1975) Logic and conversation . In P. Cole, ed. and J.L. Morgan (eds), Syntax and Semantics: Speech Acts , pp. 41–58. New York: Academic Press.

Hamilton, D.H., Katz, L.B. and Leirer, V.O. (1980) Organizational processes in impression formation . In R. Hastie, ed. , T. Ostrom, ed. , E. Ebbesen, ed. , R. Wyer, ed. , D. Hamilton, ed. , and D. Carlston (eds) Person Memory: The Cognitive Basis of Social Perception , pp. 121–153. Hillsdale, NJ: Lawrence Erlbaum Associates.

Hastie, R. and Kumar, P.A. (1979) Person memory: Personality traits as organizing principles in memory for behaviors . Journal of Personality and Social Psychology , 37, 25–38. <http://dx.doi.org/10.1037/0022-3514.37.1.25>

Hastie, R., Ostrom, T.M., Ebbesen, E.B., Wyer, R.S., Hamilton, D.L. and Carlston, D.E. (1980) Person Memory: The Cognitive Basis of Social Perception . Hillsdale, NJ: Lawrence Erlbaum Associates.

Higgins, E.T. (1996) Knowledge activation: Accessibility, applicability, and salience . In E.T. Higgins, ed. and A. Kruglanski (eds), Social Psychology: Handbook of Basic Principles , pp. 133–168. New York: Guilford.

Higgins, E.T., Rholes, W.S. and Jones, C.R. (1977) Category accessibility and impression formation . Journal of Experimental Social Psychology , 13, 141–154. <http://dx.doi.org/10.1016/S0022-1031%2877%2980007-3>

Hung, I.W.P. and Wyer, R.S. (2008) The role of implicit theories in the impact of problem-solving print advertisements . Journal of Consumer Psychology , 18, 223–235. <http://dx.doi.org/10.1016/j.jcps.2008.04.011>

Hung, I.W. and Wyer, R.S. (2009) Differences in perspective and the influence of charitable appeals: When imagining oneself as the victim is not always beneficial . Journal of Marketing Research , 46, 421–434 <http://dx.doi.org/10.1509/jmkr.46.3.421>

Jiang, Y. (2008) The role of visual imagery and visual perspective in consumer behavior. PhD dissertation , Hong Kong University of Science and Technology.

Jiang, Y. and Wyer, R. S. (2009) The role of visual perspective in information processing . *Journal of Experimental Social Psychology* , 45, 486–495. <http://dx.doi.org/10.1016/j.jesp.2008.12.006>

Kruglanski, A.W., Shah, J.Y., Fishbach, A., Friedman, R., Chun, W. and Sleeth-Keppler, D. (2002) A theory of goal systems . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 34, 331–378. San Diego: Academic Press.

Kruglanski, A.W., Thompson, E.P. and Spiegel, S. (1999) Separate or equal? Bimodal notions of persuasion and a single process ‘unimodel’ . In S. Chaiken, ed. and Y. Trope (eds), *Dual-process Theories in Social Psychology* , pp. 293–313. New York: Guilford.

Lingle, J.H. and Ostrom, T.M. (1979) Retrieval selectivity in memory-based impression judgments . *Journal of Personality and Social Psychology* , 37, 180–194. <http://dx.doi.org/10.1037/0022-3514.37.2.180>

McGuire, W.J. (1960) A syllogistic analysis of cognitive relationships . In M.J. Rosenberg, ed. and C.I. Hovland (eds), *Attitude Organization and Change* , pp. 140–162. New Haven: Yale University Press.

McGuire, W.J. (1968) Personality and susceptibility to social influence . In E. Borgatta, ed. and W. Lambert (eds), *Handbook of Personality Theory and Research* . Chicago: Rand-McNally.

Nisbett, R.E. and Ross, L. (1980) *Human Inference: Strategies and Shortcomings of Social Judgments* . Englewood Cliffs, NJ: Prentice Hall.

Petty, R.E. and Cacioppo, J.T. (1986) *Communication and Persuasion: Central and Peripheral Routes to Attitude Change* . New York: Springer-Verlag. <http://dx.doi.org/10.1007/978-1-4612-4964-1>

Radvansky, G.A., Wyer, R.S., Curiel, J.M. and Lutz, M.F. (1997) Mental models and abstract relations . *Journal of Experimental Psychology: Learning, Memory, and Cognition* , 23, 1233–1246. <http://dx.doi.org/10.1037/0278-7393.23.5.1233>

Ratcliff, R. (1978) A theory of memory retrieval . *Psychological Review* , 85, 59–108.
<http://dx.doi.org/10.1037/0033-295X.85.2.59>

Rosen, N.A. and Wyer, R.S. (1972) Some further evidence for the ‘Socratic effect’ using a subjective probability model of cognitive organization . *Journal of Personality and Social Psychology* , 24, 420–424. <http://dx.doi.org/10.1037/h0033665>

Schwarz, N. and Bohner, G. (2001) The construction of attitudes . In A. Tesser, ed. and N. Schwarz (eds), *Blackwell Handbook of Social Psychology: Intraindividual Processes* , Vol. 1. Oxford: Blackwell. <http://dx.doi.org/10.1002/9780470998519>

Shah, J.Y. and Kruglanski, A.W., (2002) Priming against your will: How goal pursuit is affected by accessible alternatives . *Journal of Experimental Social Psychology* , 38, 368–383. <http://dx.doi.org/10.1016/S0022-1031%2802%2900005-7>

Shah, J.Y., Kruglanski, A.W. and Friedman, R. (2003) Goal systems theory: Integrating the cognitive and motivational aspects of self-regulation . In S.J. Spencer, ed. and S. Fein (eds), *Motivated Social Perception: The Ontario Symposium* , 9, 247–275. Mahwah, NJ: Erlbaum.

Shen, H. and Wyer, R.S. (2008) Procedural priming and consumer judgments: Effects on the impact of positively and negatively valenced information . *Journal of Consumer Research* , 34, 121–111. <http://dx.doi.org/10.1086/523292>

Shen, H. and Wyer, R.S. (2009) Cognitive and motivational influences of past behavior on variety seeking . Unpublished manuscript, Hong Kong University of Science and Technology.

Shen, H. and Wyer, R.S. (2010) The effect of past behavior on variety seeking: Automatic and deliberative influences . *Journal of Consumer Psychology* , 20, 33–42. <http://dx.doi.org/10.1016/j.jcps.2009.07.002>

Smith, E.R. (1990) Content and process specificity in the effects of prior experiences . In T.K. Srull, ed. and R.S. Wyer (eds), *Advances in Social Cognition* , 3, 1–59. Hillsdale, NJ: Lawrence Erlbaum Associates.

Srull, T.K. (1981) Person memory: Some tests of associative storage and retrieval models . *Journal of Experimental Psychology: Human Learning and Memory* , 7, 440–463. <http://dx.doi.org/10.1037/0278-7393.7.6.440>

Srull, T.K. and Wyer, R.S. (1980) Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgments . *Journal of Personality and Social Psychology* , 38, 841–856. <http://dx.doi.org/10.1037/0022-3514.38.6.841>

Srull, T.K. and Wyer, R.S. (1983) The role of control processes and structural constraints in models of memory and social judgment . *Journal of Experimental Social Psychology* , 19, 497–521. <http://dx.doi.org/10.1016/0022-1031%2883%2990013-6>

Srull, T.K. and Wyer, R.S. (1989) Person memory and judgment . *Psychological Review* , 96, 58–63. <http://dx.doi.org/10.1037/0033-295X.96.1.58>

Strack, F. and Deutsch, R. (2004) Reflective and impulsive determinants of social behavior . *Personality and Social Psychology Review* , 8, 220–247. http://dx.doi.org/10.1207/s15327957pspr0803_1

Upshaw, H.S. (1965) The effect of variable perspectives on judgments of opinion statements for Thurstone scales: Equal-appearing intervals . *Journal of Abnormal and Social Psychology* , 64, 85–96.

Wyer, R.S. (1969) Effects of general response style on own attitude and the interpretation of attitude-relevant messages . *British Journal of Social and Clinical Psychology* , 8, 104–115. <http://dx.doi.org/10.1111/j.2044-8260.1969.tb00595.x>

Wyer, R.S. (1970) The quantitative prediction of belief and opinion change: A further test of a subjective probability model . *Journal of Personality and Social Psychology* , 16, 559–571. <http://dx.doi.org/10.1037/h0030064>

Wyer, R.S. (1974) Cognitive Organization and Change: An Information-processing Approach . Hillsdale, NJ: Erlbaum.

Wyer, R.S. (2004) Social Comprehension and Judgment: The Role of Situation Models, Narratives and Implicit Theories . Mahwah, NJ: Erlbaum.

Wyer, R.S. (2006) Three models of information processing: An evaluation and conceptual integration . *Psychological Inquiry* , 17, 185–194. http://dx.doi.org/10.1207/s15327965pli1703_4

Wyer, R.S. (2008) The role of knowledge accessibility in cognition and behavior: Implications for consumer information processing . In C. Haugvedt, ed. , F.R. Kardes, ed. , and P.M. Herr (eds), *Handbook of Consumer Psychology* , pp. 31–76. Mahwah, NJ: Erlbaum.

Wyer, R.S., Adaval, R. and Colcombe, S.J. (2002) Narrative-based representations of social knowledge: Their construction and use in comprehension, memory and judgment . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 34, 131–197. San Diego: Academic Press.

Wyer, R.S., Bodenhausen, G.V. and Srull, T.K. (1984) The cognitive representation of persons and groups and its effect on recall and recognition memory . *Journal of Experimental Social Psychology* , 20, 445–469. <http://dx.doi.org/10.1016/0022-1031%2884%2990037-4>

Wyer, R.S. and Budesheim, T.L. (1987) Person memory and judgment: The impact of information that one is told to disregard . *Journal of Personality and Social Psychology* , 53, 14–29. <http://dx.doi.org/10.1037/0022-3514.53.1.14>

Wyer, R.S., Budesheim, T.L. and Lambert, A.J. (1990) Person memory and judgment: The cognitive representation of informal conversations . *Journal of Personality and Social Psychology* , 58, 218–238. <http://dx.doi.org/10.1037/0022-3514.58.2.218>

Wyer, R.S., Budesheim, T.L., Lambert, A.J. and Swan, S. (1994) Person memory and judgment: Pragmatic influences on impressions formed in a social context . *Journal of Personality and Social Psychology* , 66, 254–267. <http://dx.doi.org/10.1037/0022-3514.66.2.254>

Wyer, R.S. and Carlston, D.E. (1979) Social Cognition, Inference and Attribution . Hillsdale, NJ: Erlbaum.

Wyer, R.S., Clore, G.L. and Isbell, L.M. (1999) Affect and information processing . In M.P. Zanna (ed.), Advances in Experimental Social Psychology , 31, 1–77. San Diego: Academic Press.

Wyer, R.S. and Collins, J.E. (1992) A theory of humor elicitation . Psychological Review , 99, 663–688. <http://dx.doi.org/10.1037/0033-295X.99.4.663>

Wyer, R.S. and Gruenfeld, D.H. (1995) Information processing in social contexts: Implications for social memory and judgment . In M.P. Zanna (ed.) Advances in Experimental Social Psychology , 27, 49–91. San Diego: Academic Press.

Wyer, R.S. and Hartwick, J. (1980) The role of information retrieval and conditional inference processes in belief formation and change . In L. Berkowitz (ed.) Advances in Experimental Social Psychology , 13, 241–284. New York: Academic Press.

Wyer, R.S., Hung, I.W. and Jiang, Y.W. (2008) Visual and verbal processing strategies in comprehension and judgment . Journal of Consumer Psychology , 18, 244–257. <http://dx.doi.org/10.1016/j.jcps.2008.09.002>

Wyer, R.S. and Martin, L.L. (1986) Person memory: The role of traits, group stereotypes, and specific behaviors in the cognitive representation of persons . Journal of Personality and Social Psychology , 50, 661–675. <http://dx.doi.org/10.1037/0022-3514.50.4.661>

Wyer, R.S. and Radvansky, G.A. (1999) The comprehension and validation of social information . Psychological Review , 106, 89–118. <http://dx.doi.org/10.1037/0033-295X.106.1.89>

Wyer, R.S. and Srull, T.K. (1980) The processing of social stimulus information: A conceptual integration . In R. Hastie, ed. , T.M. Ostrom, ed. , E.B. Ebbesen, ed. , R.S. Wyer, ed. , D.L. Hamilton, ed. , and D.E. Carlston (eds), Person Memory: The Cognitive Basis of Social Perception . Hillsdale, NJ: Erlbaum.

Wyer, R.S. and Srull, T.K. (1986) Human cognition in its social context . Psychological Review , 93, 322–359. <http://dx.doi.org/10.1037/0033-295X.93.3.322>

Wyer, R.S. and Srull, T.K. (1989) Memory and Cognition in Its Social Context . Hillsdale, NJ: Erlbaum.

Xu, A.J. and Wyer, R.S. (2007) The effect of mindsets on consumer decision strategies . Journal of Consumer Research , 34, 556–566. <http://dx.doi.org/10.1086/519293>

Xu, A.J. and Wyer, R.S. (2008) The comparative mindset: From animal comparisons to increased purchase intentions . Psychological Science , 19, 859–864. <http://dx.doi.org/10.1111/j.1467-9280.2008.02169.x>

<http://dx.doi.org/10.4135/9781446249215.n9>

Handbook of Theories of Social Psychology: Volume 1

Balance-Logic Theory

Contributors: Chester A. Insko

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Balance-Logic Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n10>

Print pages: 178-201

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n10>

[p. 178 ↓]

Chapter 9: Balance-Logic Theory

Chester A. Insko, ed.

Abstract

Following an overview of Heider's classic statement of balance theory as a description of the perceived coherence, or lack of coherence, of interpersonal relations, this chapter reviews subsequent developments including Cartwright and Harary's (1956) application of graph theory, attitude structure, two-valued logic, self-esteem, the overlap between self-consistency and hedonism, conformity, dissonance, and consideration of the tetrahedron model for generalizing the multiplicative rule beyond two-valued distinctions.

Heider's (1946, 1958) Statement of Balance Theory

Elements, Relations, and the Concept of Balance

Heider formulated balance theory as an extension of Gestalt principles to the perceived coherence, or lack of coherence, of interpersonal relations. He focuses on three types of elements and two types of relations. The elements, typically symbolized p , o , and x , are respectively the person whose experience is the focus of the theory, some other perceived person, and some perceived object or concept. Occasionally Heider also

refers to a third person as q . Regardless as to the number of persons, it is important to recognize that these are persons as perceived by p , and also that x refers to objects or concepts that are perceived by p .

Heider classifies the relations between pairs of elements as either sentiment relations or unit relations. Sentiment relations are attitudinal, or evaluative, relations that can be either positive (“liking,” “approving,” “valuing”) or negative (“disliking,” “disapproving,” “devaluating”). Positive sentiment relations are symbolized as L and negative sentiment relations as nL .

Heider (1958: 176) defines unit relations as relations that “are perceived as belonging together.” The concept of unit relations involves perceived class, or category, inclusion and, like sentiment relations, can be either positive (“similar,” “close,” “facilitates,” “belongs to”) or negative (“dissimilar,” “far,” “interferes with,” “does not belong to”). Positive unit relations are symbolized as [p. 179 ↓] U and negative unit relations as nU . Heider relates the concept of unit relations to Wertheimer's (1923) grouping principles of similarity, proximity, and common fate (or covariation).

Cartwright and Harary (1956) point out that it is important not to confuse the negation of a positive unit relation with a negative unit relation. Thus if “owns” is a positive unit relation, the opposite, negative relation is “sells,” and not “does not own.” This implies, contrary to Heider, that “does not belong to” is not a negative unit relation.

Heider (1958: 180) refers to balanced relations between elements as “a harmonious state” in which elements “fit together without stress,” and further elaborates the theory by discussing perceived homogeneity of another person, balance in dyads, and balance in the p - o - x triad.

Perceived Homogeneity of Another Person

According to Heider (1958: 183), “if several parts, or traits, or aspects of a person are considered, the tendency exists to see them all as positive, or all as negative.” This perceived homogeneity is illustrated by the well-known halo effect in the judgment of different traits in another person. Anderson (1981: 380) presents evidence for the

halo effect and concludes that “the general impression of the person acts as a causal mediator in judging specific traits of that person.”

Balance in Dyads

Restricting attention to the p , o , and x elements, there are three possible dyads: $p-x$, $p-o$, and $o-x$, but Heider (1946, 1958) only discusses the first two. The $p-x$ dyad is balanced if the unit relation and sentiment relation have the same signs, for example, “ p likes the things he made; p wants to own the things he made; p values what he is accustomed to” (1946: 108). Possible evidence for balance in the $p-x$ dyad comes from the so-called secondary reinforcement effect in which a previously neutral stimulus takes on the valence of a circumstance (e.g., electric shock or food) with which the stimulus has been repeatedly paired (cf. Miller, 1951), and Zajonc's (e.g., 1968, 2001) repeated demonstrations of preference for the familiar. An intriguing illustration of preference for the familiar is Mita et al.'s (1977) finding that participants preferred a photograph of themselves with the familiar reversed image seen in a mirror to a photograph of themselves with the less familiar image seen by others.

The $p-o$ dyad, like the $p-x$ dyad, is balanced if the unit and sentiment relation have the same signs, for example, “ p likes his children, people similar to him; p is uneasy if he has to live with people he does not like; p tends to imitate admired persons; p likes to think that loved persons are similar to him” (Heider 1946: 108). Evidence for this type of balance comes from Saegert et al.'s (1973) two experimental demonstrations in which participants expressed more liking for others when they judged tastes in the same as opposed to different booths. A manipulation of the number of trials in the same booth from 0 to 1 to 2 to 5 to 10 resulted in a corresponding increase in attraction. Insko and Wilson (1977) further found that actual interaction beyond simple contact can increase attraction. In their study, three participants, labeled A , B , and C , were seated in a triangular pattern facially oriented toward each other. Initially A and B were instructed to interact and get acquainted while C listened, and then B and C were instructed to interact and get acquainted while A listened. A subsequent assessment of attraction indicated that liking was greater between participants who interacted (A with B , B with A , B with C , C with B) than between participants who did not interact (A with C , C with A).

A further consideration not relevant to the p - x triad is that balance in the p - o triad requires reciprocation of the sentiment [p. 180 ↓] relations. As stated by Heider (1946: 108), “a balanced state exists if pLo and oLp (or $pnLo$ and $onLp$) are true at the same time.” Support for the presence of reciprocated sentiment comes from both nonexperimental and experimental studies. For example, in an early balance study, Wiest (1965) found a positive correlation of 0.74 between school-children’s reported liking for classmates and the extent to which classmates were perceived as liking them. Experimental evidence in which there is implied positive or negative evaluation of a participant by another person (e.g., Aronson and Worchel, 1966; Byrne and Griffitt, 1966; Byrne and Rhamey, 1965; Insko et al., 1973; Montoya and Insko, 2007) have consistently found evidence for reciprocated sentiment, sometimes referred to as the implied-evaluation effect.

The concept of reciprocated sentiment is, of course, not original with Heider. Shakespeare based his play *Much Ado About Nothing* on the assumption of reciprocated sentiment. Berscheid and Walster (1978) point out that this concept was the underlying assumption in Dale Carnegie’s (1937) book, *How to Win Friends and Influence People*, but that the basic idea can be traced back to a statement by the philosopher Hecato in the second century BC: “I will show you a love potion without drug or herb or any witch’s spell; if you wish to be loved, love.”

As subsequent discussion will emphasize, there are frequently additional considerations indicating that focusing on a single dyad (or triad) is an oversimplification. Two initial examples of such additional considerations occur in Heider’s discussion of the proverbs “opposites attract” and “familiarity breeds contempt.” Heider argues that if opposites do attract it is because the differences allow for the realization of some goal, and implies that for p to achieve a goal is balanced. Attraction between the sexes is an obvious example. Heider also argues that if familiarity does breed contempt it is because of too much dissimilarity. Presumably with increasing contact and familiarity there is an increasing opportunity to discover dissimilarities.

Balance in the P-O-X Triad

The *p-o-x* triad involves the perceived occurrence of three relations: *p-o*, *p-x*, and *o-x*. These relations can be either sentiment or unit relations that are either positive or negative, but Heider's theory primarily emphasizes the sign of the relation, "[I]n many cases the effects of *L* and *U* in these configurations seem to be the same" (1946: 109). On the other hand, Heider also states that "The equivalence of the *L* and *U* relations seems to be limited by the fact that often the *U* relation is weaker than the *L* relation," (1946: 111). This issue is partially addressed by Cartwright and Harary's (1983) above-described admonition to regard a negative unit relation, not as the negation of a positive unit relation, but as the opposite of a positive unit relation.

Restricting attention to just the sign of a relation, there are eight possible triads: + + +, + - - + - +, + + - , + - + - + +, - -. Following an early study by Jordan (1953), the tradition has been to refer to the first sign as relating to the *p-o* (I to other) relation, the second sign as relating to the *p-x* (I to object) relation, and the third sign as relating to the perceived *o-x* (other to object) relation.

Heider considers four of the eight triads (+ + +, + - , - + -, - +) to be balanced. These are triads in which all of the relations are positive, or two of the relations are negative and one is positive. An example of a + + + triad is: "*p* likes *o*, likes the book *x*, and perceives that *o* wrote the book *x*." An example of a - + triad is: "*p* dislikes *o*, dislikes the book *x*, and perceives that *o* wrote the book *x*."

Heider considers three of the remaining relations in which there is a single negative sign (+ + -, + - +, - + +) as imbalanced, for example "*p* likes the author of a disliked book." The remaining – triad, however, is considered ambiguous: "[T]he case with three negative signs does not constitute a good psychological balance, since it is too indetermined" (1946: 110). Following Cartwright and Harary (1956), subsequent researchers have universally considered the – triads as imbalanced.

[p. 181 ↓] As pointed out above, it is frequently an oversimplification to consider just a single dyad or triad because *p*'s experience may include multiple dyads and triads. For example, a critic may state: "Just because I disagree with my friend about some

issue it is foolish to assume that I am therefore going to dislike my friend." Indeed it may be. Note that while Byrne (1971) is well known for advocating the effect of agreement or similarity on attraction, even he does not claim an effect for the similarity of a single attitude, but rather an effect for the proportion of similar attitudes. Research has further demonstrated that the effect of the proportion of similar attitudes linearly increases as the total number of attitudes increases from 3 to 6 to 12 to 24 (Wetzel and Insko, 1974).

An additional example of the importance of considering more than one dyad or triad is the so-called "love triangle" in which two male friends are attracted to the same female. This is a + + + triad that common sense regards as unstable. Heider's discussion of this situation is succinct and to the point: "*p* does not want his girl friend *o* to fall in love with his boy friend *q* because *oLq* in this case implies *onLp*, which conflicts with *pLo*" (1946: 110).

Reactions to Imbalance

Heider (1958) discusses reactions to imbalance by describing a simple study in which participants were asked to write how they thought a hypothetical person, Bob, would react to an imbalanced situation:

Bob thinks Jim is very stupid and a first class bore. One day Bob reads some poetry he likes so well that he takes the trouble to track down the author in order to shake his hand. He finds that Jim wrote the poems (1958: 176).

Heider coded the written reactions into five categories, and noted the frequency of each. First, 46 percent described Bob as feeling less negative about Jim; for example, "He grudgingly changes his mind about Jim" (1958: 186). Second, 29 percent described Bob as feeling less positive about the poetry; for example, "He decides the poems are lousy" (1958:186). Third, 5 percent described Bob as doubting that Jim wrote the poems; for example, "Bob would probably question Jim's authorship of the poems" (1958: 176–7). Fourth, 2 percent described Bob as differentiating or separating, two aspects of Jim; for example, "He thinks Jim is smart in some lines but dumb in others" (1958: 177). Fifth and finally, 18 percent did not describe Bob as resolving the

imbalance but some were aware of the tension; for example, “Bob is confused and does not know what to do. He finally briefly mentions liking of the poems to Jim without much warmth” (1958: 177).

Heider interprets the fifth reaction as illustrating the important point that while imbalance will create tension, that tension will not always be resolved. In this study, of course, most of the participants did describe Bob as resolving the imbalance. The first three reactions involve simple changes in one of the relations in the *p-o-x* triad. The fourth, or differentiation, reaction is less simple.

Differentiation involves a reconceptualization of a single element as two elements. If the element is an *x*, that can sometimes be easily accomplished. For example, two friends who disagree about the value of foreign aid might differentiate foreign aid into military aid, which they both dislike, from humanitarian aid, which they both like. An experiment by Stroebe et al. (1970) suggests that differentiation can also occur if the element is a person.

In the Stroebe et al. (1970) experiment participants were given information about a person, Dr. M., indicating that he was an expert or inexpert scientist and a nice or awful person. Whether Dr. M. was an expert or inexpert scientist affected evaluation of a theory with which he was associated but not of his wife. On the other hand, crosscutting information that Dr. M. was a nice or awful person affected evaluation of his wife but not of the theory. The results thus indicated that [p. 182 ↓] the participants differentiated Dr. M. the scientist from Dr. M. the person.

Abelson (1959) has suggested two reactions to imbalance other than the ones listed by Heider. One of these is bolstering, or strengthening, of one of the relations producing imbalance. In terms of the above example, Bob might recall past episodes in which Jim had behaved in a boring manner. The other is transcendence, or the development of a theory accounting for the inconsistency between or among inconsistent elements. For example, Bob might develop a point of view indicating that it is human nature for poetry to be written by boring people.

Osgood and Tannenbaum (1955), Abelson and Rosenberg (1958), and McGuire (1960) have made other suggestions regarding reactions to imbalance or inconsistency.

Addressing themselves to change in a relation, Osgood and Tannenbaum (1955) proposed that change is proportional to the degree of polarization in the relation. Addressing themselves to a network of relations, Abelson and Rosenberg (1958) proposed a least-effort principle according to which the change that occurs will be the one that requires the fewest number of additional changes. Finally, also relevant to a network of relations, McGuire (1960) proposed a cognitive inertia principle according to which changes in more remote relations will require more time than changes in relations closest to an encountered inconsistency.

Cartwright and Harary's Multiplicative Rule

Cartwright and Harary (1956) address balance theory's elements and relations with a graph theory representation of points and lines. Positive relations are indicated with solid lines and negative relations with dashed lines. Such an approach allows for the easy representation of sets of relations involving more elements, or points, than dyads or triads. A closed loop of points and lines is referred to as a *cycle*. In a situation in which it becomes important to distinguish the direction of a relation, for example the difference between *p* to *o* and *o* to *p*, the lines are given arrowheads and the interconnected points are referred to as a *semicycle*. Thus the *p* to *o* and *o* to *p* dyad is a semicycle, as is the *p* to *o*, *p* to *x*, *o* to *x* triad.

A cycle or semicycle is balanced only if the product of the signs is positive. Thus, for example, *p-o-x* triads, or semicycles, that contain three positive signs or one positive sign and two negative signs are balanced, and *semicycles* that contain one negative sign and two positive signs or three negative signs are imbalanced. Note that this multiplicative rule applies equally well to Heider's *p-x* dyad, or semicycle, that is balanced if the sentiment relation and unit relation have the same sign, and also to the *p-o* dyad, or semicycle, that has the additional consideration that the *p* to *o* relation and the *o* to *p* relation have the same sign. If a set of positive and negative lines, referred to as a signed graph or an *s*-graph, contains more than one cycle, or semicycle, that graph is balanced if and only if all cycles, or semicycles, are balanced.

Cartwright and Harary state a theorem regarding balance in an *s*-graph: "An *s*-graph is balanced if and only if its points can be separated into two mutually exclusive subsets

such that each positive line joins two points of the same subset and each negative line joins two points from different subsets" (1956: 286). This theorem describes what is sometimes referred to as a "black and white" attitude in which, for example, p perceives positive sentiment relations among the ingroup members and negative sentiment relations with the similarly perceived out-group members. The theorem also describes one interpretation of the Judeo-Christian conception of heaven and hell.

Cartwright and Harary's consideration of multiple linkages points to new and important topics beyond dyads and triads, for example, attitude structures and conformity effects discussed below. Cartwright and Harary's discussion also raises an issue [p. 183 ↓] regarding the utility of consistent thought. It is clear, as Cialdini (2009) points out, that without a tendency to think consistently "our lives would be difficult, erratic, and disjointed" (2009: 53). However, Cialdini also correctly indicates that the tendency to think consistently can create a "foolish fortress" (2009: 54) in which premature closure leads to disregard of important facts and considerations. The mixed utility of consistency is captured in Ralph Waldo Emerson's frequently quoted statement: "A foolish consistency is the hobgoblin of little minds, adored by little statesmen and philosophers and divines." Consider, for example, consistent application of the "Reaganomics" principle that "government is the problem."

Attitude Structures and Cognitive Responses

While Heider developed balance theory primarily in the context of interpersonal relations, Peak (1955), Rosenberg (1956, 1960), and Fishbein took a similar, consistency approach to attitudes that were related with linking beliefs. Rosenberg refers to a single attitude object with linking beliefs to other attitude objects, or values, as having a structure. He illustrates the basic idea by describing a doctor who negatively evaluates socialized medicine. The doctor believes that socialized medicine interferes with high medical standards and positively evaluates high medical standards. Further, the doctor also believes that socialized medicine facilitates government control and negatively evaluates government control. Each of these two examples (socialized medicine interferes with high medical standards, and socialized medicine facilitates

government control) is referred to as a cognitive band. Rosenberg sees an attitude structure as consisting of a series of such cognitive bands. The linking beliefs in each cognitive band of particular interest to Rosenberg were instrumental, or facilitates-interferes-with, cognitions. In Heider's language these are positive or negative unit relations.

Some readers may question what happened to *p*. The person, *p*, is the doctor. Thus a cognitive band could be stated as a *p-x-y* triad; the doctor, *p*, dislikes socialized medicine (-), dislikes government control (-), and believes that socialized medicine facilitates government control (+).

Rosenberg assumes that attitude structures tend to maintain a stable homeostatic, or consistent, state. One procedure that Rosenberg followed to test this theory involved a number of steps. First, participants responded to a +3 (strongly facilitates) to -3 (strongly interferes with) scale relating a central attitude object to a number of other attitude objects, or values. Second, participants responded to a +3 (very favorable) to -3 (very unfavorable) scale for each of the linked values. Third, the belief and evaluative ratings within each cognitive band were multiplied. Fourth, the products of the belief by evaluative ratings within each cognitive band were summed across bands to give an index expected to predict attitude toward the central attitude object.

Rosenberg (1956, 1960) found that the sum of products across cognitive bands significantly predicted attitude toward free speech for communists, and segregated African-American housing, two topical issues of the time. Other investigators who used variations of Rosenberg's procedure found similar results. The earliest of these studies by Woodruff and DiVesta (1948) was concerned with attitude toward the abolishment of fraternities and sororities. A subsequent study by Fishbein (1963) was concerned with attitude toward African Americans, and a study by Insko et al. (1970) was concerned with attitude toward the use of birth control. These studies found evidence for consistent attitude structures.

One approach to developing an attitude-structure questionnaire involves the conducting of prior interviews with a sample of respondents who are probed for beliefs related to the attitude issue in question. [p. 184 ↓] An alternative approach involves the use of a thought-listing technique originally developed as a measure of cognitive responses

to a persuasive message (Brock, 1967; Greenwald 1967, 1968). The technique requires that participants list their thoughts, one thought per line. Brock had judges code the listed thoughts for counterarguments to the point of view advocated in a communication, and then created an index that was the simple sum of the number of such counterarguments. Greenwald had the participants themselves categorize their thoughts as favorable or unfavorable to the point of view advocated in the communication, and then created an index that was the difference between the number of favorable and unfavorable thoughts. Both Brock and Greenwald found that their indices were related to the participants' postcommunication attitude. Greenwald, for example, found a positive correlation of 0.65 between his index and postcommunication attitude.

Even though the possible relevance of listed thoughts to cognitive bands was pointed out (Insko, 1981), the relevance was generally not recognized. Research with the listed-thought technique has revealed that some issue-relevant thoughts are explicit cognitive bands; for example, an explicit counterargument, and some others can be easily interpreted as implicit cognitive bands. For example, if following a communication advocating the building of nuclear power-generating plants, a respondent lists the thought "problem of radioactive waste," there is an implicit reference to a consistent cognitive band: "Nuclear power generating plants are bad (-) because they generate (+) dangerous radioactive waste (-)."

Two-Valued Balance and Two-Valued Logic

It is clear that Heider regarded balance theory as an application of Gestalt principles rather than logic. For example, in his posthumously published notes (Benesh-Weiner, 1988), Heider writes: "For balance: there is a gestalt quality (pleasantness, fittingness) correlated to a 'structure.' The structure is a relation between the relations, between the parts ..." (1988: 52).

In an influential article Abelson and Rosenberg (1958) distinguished between logic and psycho-logic (their term for balance). However, two logicians, Runkel and Peizer (1968) pointed out that if relations are restricted to two-valued, plus or minus, distinctions, the multiplicative rule is perfectly mapped by two-valued logic. Quoting directly: "Once it is

realized that at most two categories are available for assigning perceptually associated elements, the practical distinction between psycho-logic and ordinary logic becomes superfluous from the point of view of the behavioral scientist" (1968: 61).¹

As illustrated by the familiar syllogism in which Socrates either is or is not mortal, it is important to recognize that deductive logic makes only two-valued distinctions, and thus in order for logic to map the implications of balance theory, balance-imbalance must also be restricted to two-valued distinctions. Working through the eight possible *p-o-x* triads makes this mapping fairly obvious. For example, restricting the similar-dissimilar dimension to the two values of same and different, it follows logically that if *a* is the same as *b*, and *b* is different from *c*, then *a* is different from *c*, consistent with the multiplicative rule (+ -).

Seemingly deductive logic's two-valued assumption is not always understood. A well-known social psychologist once commented: "5 ≠ 6 and 6 ≠ 7 does not imply that 5 = 7." The problem here is that several degrees of inequality were assumed. Note that if only the values of equal and unequal are allowed, *a* ≠ *b* and *b* ≠ *c* does imply that *a* = *c*.

Abelson appears not to have been convinced by Runkel and Peizer (1968) because at a later time he critiqued balance theory as due to its historical dependence on a "weak ... linkage drawn between configurations of stimulus elements and configurations of [p. 185 ↓] cognitive elements" (1983: 411). Harary (1983) replied to Abelson that the basis of balance theory is really Boolean algebra – the logical calculus of truth values that results in the same conclusions as the semantic approach to logic.

There appears to have been a general lack of appreciation of the logical basis of the multiplicative rule. Newcomb (1968), for example, developed a version of balance theory in which balance was restricted to triads, or semicycles, in which the *p* to *o* relation is positive. This, of course, departs from the multiplicative rule and logic. But are participants responsive to balance in semicycles in which the *p* to *o* relation is negative? Aronson and Cope (1968) found that just as participants inferred that my friend's friend is my friend, they also inferred that my enemy's enemy is my enemy (see also Gawronski and Walther, 2008, Experiment 2).

Henle on Logical Thinking

The fact that the multiplicative rule is mapped by logic raises a question regarding whether thinking is logical. Some social psychologists have assumed that thinking is not logical. For example, Bruner et al. wrote: “Much of human reasoning is supported by a kind of thematic process rather than by an abstract logic” (1956: 104).

Henle (1962) points out that whether or not thought is logical has been an issue of dispute among philosophers. An older generation of philosophers, for example, John Stuart Mill (1874), regarded logic as descriptive of the laws of thought. However, more recent philosophers – for example, M.R. Cohen – have argued otherwise: “That the laws of logic are not the universal laws according to which we do actually think is conclusively shown not only by the most elementary observation of introspection, but by the very existence of fallacies” (Cohen 1944: 2–3).

Henle notes that this change in philosophical opinion appears to have been based more on a cultural shift than on the discovery of new evidence, and that the older philosophers were also aware of error. For example, Mill wrote that a person “... has it almost always in his power to make the syllogism good by introducing a false premise; and hence it is scarcely ever possible decidedly to affirm that any argument involves a bad syllogism” (1874: 560).

The fact that the philosophical disagreement related to the interpretation of error provided Henle with a justification for collecting open-ended data on participants' reports of their thoughts. Psychology graduate students were presented with a series of syllogisms in the context of everyday discussions, and then asked to state whether the conclusion of each syllogism followed and to describe their reasoning. As expected errors did occur, but the stated reasons for the false judgments could be categorized as: (1) failure to accept the requested logical task; (2) alteration of the meaning of a premise; (3) omission of one or more premises; and (4) adding a premise.

Henle gives two interesting examples of the prevalence and importance of syllogisms in everyday life, syllogisms that Aristotle referred to as “practical syllogisms”:

It is too far to walk to the Public Library; I must take a subway or bus.
The fifth Avenue bus passes the Public Library. I do not want to wear
the same dress two days in succession. I wore this dress yesterday; so
I do not want to wear this dress today (1945: 374).

Henle argues that without reliance on the practical syllogism it is not obvious that people could cope with the “ordinary tasks of life ... understand each other, follow one another's thinking, reach common decisions, and work together” (1945: 374).

Despite such examples of explicit reasoning, Henle quotes, with apparent agreement, Aristotle's assertion that the reasoning in the practical syllogism may be implicit:

The mind does not stop and consider at all one of the two premises, namely, the obvious one; for example, if walking is good for a man, one does [p. 186 ↓] not waste time over the premise “I am myself a man.” Hence such things as we do without calculation, we do quickly” (1945: 701).

Although Henle does not discuss balance theory, it is arguable that syllogisms can be recast as semicycles. For Henle's two above examples we have balanced + - semicycles: “Taking the fifth avenue bus (+) avoids (-) having to walk too far to the Public Library (-),” and “Wearing a different dress today (+) avoids (-) wearing the same dress two days in a row (-).” Quite possibly human thought more typically takes the form of semicycles than syllogisms. Perhaps this is why Bruner et al. (1956) made the above statement that human reasoning is more “thematic” than logical.

The Assumption of High Self-Regard

One of the least developed aspects of balance theory relates to the self. Heider (1946) does indicate that his discussion of dyads and triads is based on an assumption of high self-regard:

High self regard of p can be expressed by pLp , low self regard by $pnLp$ (although the two p 's in this expression are not strictly equivalent). All of the examples so far considered presupposed pLp (1946: 111).

In addition to an unexplored assumption of high self-regard, Heider also seems to assume the existence of two related, but different, selves. Wiest (1965) makes a similar assumption when he argues that if p has a positive unit relation with the self-concept, or s , in order for balance to occur there should also be a positive sentiment relation with the self-concept, pLs .

Although not explicitly stated, both Heider and Wiest appear to agree with James' (1890) classic distinction between the self as "I" and the self as "Me," or the self as subject and the self as object. The self as subject is the self that is assumed to direct attention to perceptions and thoughts and to multitask with a resulting increase in knowledge. The self as object, sometimes referred to as the self-concept, is the knowledge that is acquired regarding the person. Leary and Tangney (2003) relate the distinction between the two self-processes to the distinction between attention and cognition.

There is evidence that people tend to have high self-esteem, or high self-regard. This evidence appears to relate more obviously to the self as object than to the self as subject, but given Wiest's (1965) assumption of a positive unit relation between the two selves, they should be similarly evaluated. Support for unrealistically high self-regard has been documented by Taylor and Brown (1988), by Gilovich (1991), and by Alicke and associate's evidence for the "better than average" effect (e.g., Alicke, 1985). Alicke's evidence initially came from a demonstrated tendency for undergraduates to rate desirable traits as more characteristic of themselves than the average college student, and undesirable traits as less characteristic of themselves than the average college student – particularly if the traits were considered controllable.

Baumeister et al. (1989) obtained further evidence for the prevalence of high self-esteem. They found that for 20 different studies and 12 different self-esteem scales the scores were clustered toward the upper half of the possible range with the mean or median well above the mathematical midpoint. Baumeister et al. (1989) indicated that "low" self-esteem participants do not characterize themselves as possessing

strongly negative traits, but tend to select the neutral-middle of most negative to positive dimensions with “somewhat” or “sometimes” or “average” responses.

Consistent with the above evidence, a cross-national study by Schmitt and Allik (2005) obtained striking support for the prevalence of high self-esteem. In this study the widely used Rosenberg (1979) Self-Esteem Scale (RSES) was translated into 28 different languages and administered to participants in 53 different nations. The results [p. 187 ↓] indicated that, although there was cross-national variation, “All nations scored above the theoretical midpoint of the RSES, indicating generally positive self-evaluation may be culturally universal” (Schmitt and Allik, 2005: 623).²

The assumption of high self-esteem allows for balance interpretations of various effects, such as effects in the conformity and dissonance literatures discussed below. The following discussion will be limited to a single example relating to the so-called similarity-attraction effect.

Heider (1946, 1958) clearly assumes that if *p* and *o* are similar *p* should be attracted to *o*, and there is evidence (Byrne, 1971) that in many situations this generalization holds for attitudes and values. However, for traits the situation is less clear. Thus, for example, an introvert may or may not be more attracted to introverts than to extroverts. How can this apparent difference between attitudes and traits be explained? There is evidence that attraction is more closely related to similarity to ideal self than to similarity to actual self (LaPrell et al., 1990, 1991; Wetzel and Insko, 1974). Note that since in most instances attitudes are more readily changeable than are traits, actual and ideal attitudes are more likely confounded than are actual and ideal traits. Thus it follows that attraction should be more closely related to similar attitudes than to similar traits. This interpretation is consistent with Alicke's (1985) above described evidence that the “better than average” effect held more obviously for personal traits that were controllable.

Logic and Hedonism

There is a conceptually interesting overlap between logic and a hedonistic, or reward-cost, perspective that is captured by the common reference to “rational economic

decisions.” The implicit assumption of this statement is that some decisions are both hedonistic and rational in that they allow for the maximization of outcomes by following the logical implication of behaving consistently with a positive self. Note that, given high self-regard, it is both balanced and hedonic for a positive self (+) to receive (+) a reward (+) or for a positive self (+) to avoid (-) a cost (-), and that it is both imbalanced and antihedonic for a positive self (+) to avoid (-) a reward (+) or for a positive self (+) to receive (+) a cost. The overlap between logic and hedonism stands in contrast to the frequent assumption, for example McGuire (1960), that wishful thinking and logical thinking are fundamentally different. The overlap between logic and hedonism suggests two issues that will be briefly considered. One of these relates to testability and the other to the evolutionary priority of logic and hedonism.

Testability

A classic issue associated with hedonism relates to its testability. To use a contemporary example, is the suicidal behavior of terrorists inconsistent with hedonism or rather an indication that terrorists are concerned with increasing the rewards of an afterlife by harming infidels? Because the hedonistic perspective does not generally specify for every circumstance what will be regarded as a reward or cost, the theory is in a general sense not testable. This issue, however, is easily misunderstood. Even though the general theory is not testable, what is testable are specific predictions regarding the rewards or costs that will control behavior for particular people in particular situations.

Given the conceptual overlap between logic and hedonism, it follows that the testability issue applies to the behavioral implications of both perspectives. Recall the above-described argument of Henle's (1962) that the apparent occurrence of logical errors could be due to a variety of causes, such as the omission of a premise, and her quote of John Stuart Mill's assertion that “...it is scarcely ever possible decidedly to affirm that any argument involves a bad syllogism” (Mill 1874: 560).

[p. 188 ↓]

Evolutionary Priority

Given the overlap between consistency with a positive self and hedonism, there is a fundamental question regarding which of these processes played a crucial role in evolutionary development? Clearly an organism (or its genes) that did not seek rewards (such as food) or avoid costs (such as predators) would be selected against. But does such behavior flow from consistency with a positive self, or hedonism, or something more primitive than either? Actually, it is difficult to imagine anything more primitive than a simple, consistency algorithm.

Note, however, that the postulation of such an algorithm seemingly requires a positive self, and it is implausible that simple one-celled organisms possess a self. Elsewhere (Insko, 1999: 133) I have speculated that the problem could be solved by assuming the emergence of a primal-positive sign that made survival logical. While the postulation of such a primitive process is highly speculative, it is arguable that without the emergence of such a primal sign, survival would not have been possible. Note also that the postulation of a primal positive sign is no more “off the wall” than an assumption that one-celled organisms are motivated by a desire to receive rewards and avoid costs.

If the survival of primitive life required a primal-positive sign, there is the further interesting possibility that through additional evolutionary development the primal-positive sign was elaborated into a more complex self. Although they do not postulate a primal-positive sign, such a possibility is consistent with Sedikides and Skowronski's (1997) argument that rudimentary forms of the self exist in simpler, nonhuman animals.

Studies of Hypothetical *P-O-X* Semicycles

Beginning with Jordan (1953), a student of Heider's, numerous studies were conducted in which participants rated the pleasantness of each of the eight possible semicycles generated by positive or negative *p* to *o*, *p* to *x*, and *o* to *x* relations (+ + +, + -, - + -, - +, + + -, + --, - + +, --). Researchers have generally followed Jordan's precedent of letting the first sign refer to the *p* to *o* relation, the second to the *p* to *x* relation, and the third

to o to x relation. As presented to participants “I” was substituted for “ p ” as in “I like O , I like X , and O likes X .”

Jordan found that mean pleasantness-unpleasantness ratings on a scale in which low numbers indicated pleasantness were as follows: 26.2 (+ + +), 39.5 (+ - -), 55.3 (- + -), 62.4 (- - +), 57.0 (+ + -), 58.2 (+ - +), 54.8 (- + +), 58.4 (-). These data indicate an approximate tendency for the first two semicycles (+ + + and + -) to differ from the remaining six. The first two semicycles are semicycles in which p likes o and agrees with o by liking or disliking x . According to the multiplicative rule these two semicycles are balanced. However, the two further balanced semicycles, in which p dislikes o and disagrees with o (- + - and - +), were not rated as particularly pleasant.

Zajonc (1968) summarized the literature on hypothetical triads by calculating attraction, agreement, and balance indices for each study. Each index was a ratio. For example, the attraction index was the ratio of the four semicycles in which p likes o to the four semicycles in which p dislikes o . A more familiar approach would have been to view the semicycles as the eight cells generated by a three-factor analysis of variance design in which the factors were p to o , p to x , and o to x . From this perspective the attraction index, or effect, is a main effect for the p to o factor such that semicycles in which p likes o were rated as more pleasant than semicycles in which p dislikes o . The agreement effect is a double interaction of the p to x and o to x factors such that the semicycles in which p and o agree by both liking or both disliking x were rated as more pleasant than the semicycles in which p and o disagree in their liking or disliking of x . And finally, the balance effect is the triple interaction among all three [p. 189 ↓] factors such that the balanced semicycles in which p likes and agrees with o or dislikes and disagrees with o were rated as more pleasant than the imbalanced semicycles in which p dislikes and agrees with o or likes and disagrees with o .

Across all of the reviewed studies, Zajonc (1968) found that “the results are in favor of the agreement variable, with balance holding a close second, and attraction a decided third place” (1968: 347). Note that the tendency of the agreement effect to be descriptively larger than the attraction effect indicates that over all studies Jordan's (1953) two versus six pattern did not exactly replicate because in the four cells in which attraction and agreement were in opposition (+ + -, + - +, - + +, -), agreement was rated more pleasant than attraction.

The results for Zajonc's three indices led him to conclude that: "This rough summary is damaging to the balanced principle" (1968: 347). Indeed it is the case that the manipulated *p-o-x* semicycle does not adequately account for the results. However, as was observed above, it is frequently an oversimplification to assume that the only relevant semicycle is the semicycle that is observed or manipulated. The argument here parallels Henle's (1962) above-described argument that the occurrence of an apparent error in the judgment of a syllogism does not necessarily indicate that illogical thinking has occurred because the participant may, for example, add one or more additional premises. In the present context the argument is that participants may assume additional relations that create one or more additional semicycles. Beginning with a suggestion by Aderman (1969) regarding contact, these possibilities are reviewed below.³

The Assumption of P-O Contact

Aderman (1969) suggested that when the typical participant considers a hypothetical *p-o-x* semicycle, he or she tends to assume contact between *p* and *o*. Although Aderman did not explicitly develop the implications of this insight, a positive unit relation between *p* and *o* would create two additional semicycles in addition to the manipulated semicycle. These are an attraction-contact semicycle that is balanced when *p* likes and has contact with *o* and an agreement-contact semicycle that is balanced when *p* agrees with and has contact with *o*. Consideration of all three semicycles indicates that the only cells in which all semicycles are balanced are the two + + + and + - - cells that Jordan (1953) found to be more pleasant than the remaining six cells. In each of the remaining six cells only one of the three semicycles is balanced. For example in the - + + cell the agreement semicycle is balanced but the attraction semicycle and the manipulated semicycle are imbalanced.

An initial question regarding the contact hypothesis is whether participants who rate the hypothetical situations do, in fact, assume contact between themselves and the other person. Research with postexperimental questionnaires (Insko, 1984) has found that 72 to 91 percent of the participants indicated that they did assume such contact when making their ratings. Further experimental studies were conducted in which ratings were

made on five scales (pleasantness, harmony, expectancy, consistency, stability) and participants were instructed to make certain assumptions regarding contact or, in a standard condition, were given no instructions regarding contact. The results of these studies were at least somewhat supportive of the contact hypotheses. Most notably, assumed breaking *p* and *o* contact reversed the direction of the agreement effect in three of four studies and reversed the direction of the attraction effect in all four studies.

While the results, particularly the breaking-contact results, imply that assumed contact does play a role in accounting for attraction and agreement effects, clearly more than contact is involved. First, and perhaps most obvious, even when participants were instructed to assume no contact between *p* and *o*, the agreement effect was significant in [p. 190 ↓] three of the four experiments and the attraction effect was significant in one of the four experiments. Second, the contact interpretation cannot account for the fact that the agreement effect has been routinely found to be descriptively larger than the attraction effect. Third, a comparison among the five different rating scales indicated that the attraction and agreement effects were larger on the affective scales (pleasantness, harmony) than on the more cognitive scales (consistency, expectancy, stability). There is no obvious way in which the attraction-contact and agreement-contact semicycles can account for this difference.

Assumed O to P Sentiment and the Concern with Being Liked

The typical study of *p-o-x* semicycles manipulated the *p* to *o* sentiment relation but not the *o* to *p* sentiment relation. However, there should be a balance tendency to assume reciprocal *o* to *p* sentiment in the *p-o* dyad, or semicycle, and also to assume agreement-consistent *o* to *p* sentiment in the *o* to *p*, *o* to *x*, *p* to *x* triad, or semicycle. Thus the balanced-implied assumption of *o* to *p* sentiment creates additional semicycles in addition to the manipulated semicycle. Furthermore, research has indicated that when participants are asked to rate *o* to *p* liking they report more such liking when *p* likes *o* than when *p* dislikes *o* and when *p* agrees with *o* than when *p* disagrees with *o* (Insko, 1984). However, an additional semicycle is still implied.

The additional semicycle is a semicycle that includes the self as an element and that assumes the transformation of the o to p sentiment relation into an element, being liked or being disliked. Note that it is balanced for a positive self (+) to be (+) liked (+), and that it is imbalanced for a positive self (+) to be (+) disliked (-), and that this difference suggests a concern with being liked.

The postulated concern with being liked agrees with the assumption of sociometry theory that the perception that one is liked by others leads to more positive self-evaluation. Although not typically interpreted as a consistency effect, there is an abundance of experimental evidence (e.g., Leary and Baumeister, 2000; Leary et al., 2003), and also nonexperimental evidence (e.g., Srivastava and Beer, 2005) in agreement with this assumption.

In the context of research with hypothetical situations, the self-relevant semicycle implying a concern with being liked is particularly important because it suggests a reason why attraction and agreement effects should be larger with ratings on affective scales than with ratings on cognitive scales. Note that while consistency may or may not relate to the self, affect exists only in relation to a self or organism. Thus an affective rating should be particularly sensitive to semicycles including the self as an element.

The research literature (summarized in Insko, 1984) includes two experiments testing the above reasoning. One experiment included a condition in which o to p liking or disliking was consistent with agreement or disagreement, and the other included a condition in which o to p liking or disliking was consistent with p to o liking or disliking. The first experiment found that the explicit statement of o to p liking consistent with agreement increased the agreement effect relative to a standard condition, and did so primarily on the affective scales. The second experiment found that the explicit statement of o to p liking consistent with p to o liking increased the attraction effect and also did so primarily on the affective scales. Thus the assumption of o to p sentiment provides a possible explanation of the agreement and attraction effects and also that such effects are larger on affective than cognitive scales.

Assumption That O is Right and the Concern with Being Right

Although the concern with being liked provides a possible explanation as to why [p. 191 ↓] attraction and agreement effects should be larger on affective than cognitive scales, it does not account for the fact that agreement effects have repeatedly been found to be larger than attraction effects. However, social-comparison theory (Festinger, 1950, 1954a, 1954b) suggests a consideration that could create an additional semicycle relating just to agreement, and thus accounting for the larger size of the agreement effect.

According to Festinger (1950, 1954a, 1954b), people have a drive to hold correct opinions. The drive can sometimes be satisfied through simple observation (i.e., by referring to physical reality), but when this cannot be done, as for example with political, religious, or moral opinions, the drive is satisfied by agreeing with others (i.e., by referring to social reality).

As applied to hypothetical semicycles, the implication is that *o* is right about *x* and that by agreeing with *o*, *p* is also right: *o* is (+) right, *p* agrees with (+) *o*, *p* is (+) right. But when considering hypothetical *p-o-x* situations, do participants assume that the other person has some basis for his or her evaluation of *x*? A single study found that one-third did not, but that two-thirds did (see Insko, 1984).

For participants who do assume that the other person has a basis for his or her evaluation of *x*, there is the possibility of a self-relevant semicycle. Note that it is balanced for a positive self (+) to be (+) right (+), and that it is imbalanced for a positive self (+) to be (+) wrong (-), and that this difference suggests a concern with being right.

A single study that manipulated *o*'s knowledge of *x* altered the magnitude of the agreement effect and thus provided an explanation as to why agreement effects are larger than attraction effects. However, the increased magnitude of the agreement effect was not larger on the affective than the cognitive scales and thus suggests that if the semicycle is hypothetical there is little or no salience of the self-relevant concern with

being right. The following section will review some evidence that in nonhypothetical situations the concern with being right can play a role.

Conformity

Normative and Informational Social Influence

Deutsch and Gerard (1955) are widely referenced for interpreting conformity as due to two different factors: normative social influence and informational social influence. They define normative influence somewhat abstractly as “influence to conform to the positive expectations of another,” and they define informational influence as “influence to accept information obtained from another as *evidence about reality*” (1955: 629).

The Concerns with Being Liked and Being Right

Although Deutsch and Gerard (1955) did not explicitly relate normative and informational social influence to the self, there is an obvious similarity between normative social influence and self-relevant concern with being liked (or not disliked), and between informational social influence and self-relevant concern with being right (or not being wrong).

To examine the role of these two concerns in a nonhypothetical context Insko et al. (1983); and Insko et al. (1985) conducted two different experiments. Both experiments followed the Asch (1952) procedure of having a series of confederates and a single participant seated in a row of chairs, with the actual participant seated next to last. Unlike in Asch's research the judgments were not of line lengths but of colors, for example, whether the blue-green in the middle was more like the blue on the left or the green on the right. Confederates responded incorrectly on a subset of the trials.

Both experiments included two factors. One factor related to whether the participants [p. 192 ↓] responded publicly or privately. In the private condition the participants wrote their judgments on booklets that were to be dropped in a box with other unsigned booklets and “stirred” by the experimenter. The other factor related to whether the correct judgments could be determined, or not determined, by the readings of a color spectrometer.

In the determined condition, but not the undetermined condition, it was explained that the spectrometer readings would be available at the end of the experiment. If the possibility of being proven right or wrong created a greater concern with being right and if, as social-comparison theory implies, group judgments were regarded as a source of information, it follows that there should have been more conformity in the determined than the undetermined condition.

The results revealed significant main effects for both factors. The public-private main effect revealed greater conformity in the public condition than in the private condition. The determined versus undetermined main effect revealed greater conformity in the determined condition than in the undetermined condition. These main effects are consistent with both the concern with being liked and the concern with being right.

Baumeister (1982) has argued that many social phenomena, including conformity, can be explained as a manifestation of concern with self-presentation. He pointed out that “the most common procedure for testing for self-presentational motives is by comparing two situations that are identical in all respects except that some circumstance is public in one situation but private in the other” (1982: 4), and he emphasized that in order for a circumstance to be truly private, the participant must believe that the experimenter is unaware of how he or she performed.

In the private condition of the two Insko et al. (1983, 1985) experiments, participants anticipated that their booklets containing unsigned color judgments were to be dropped in a box containing other unsigned booklets and then “stirred” by the experimenter. In view of the private nature of the private condition it is important to note that the interaction between the determined versus undetermined factor and the private versus public factor was not significant in either experiment. If the effect for the determined versus undetermined factor had been due to a concern with self-presentation, there

should have been an interaction with the private versus public factor such that the difference between the determined and undetermined means was greater in the public condition. The nonsignificance of the interaction does not support an interpretation that the concern with being right was associated with, or dependent on, a self-presentational concern.

Although the evidence is inconsistent with the possibility that the concern with being right was dependent on a self-presentational concern, it is still possible that the public-private main effect was due, not just to an intrinsic concern with one's self-worth, but to appearing worthy to others. Stated differently, the greater conformity in public than private situations may have been due to a concern with "looking good," to a concern with "being good," or to both.

A Possible Role for Contact

The above-described studies of hypothetical *p-o-x* semicycles provided some evidence that assumed *p* to *o* contact played at least a minor role in producing agreement effects. Circumstantial evidence that contact may also be important in nonhypothetical situations comes from the fact that in the two Insko et al. (1983, 1985) experiments the number of conformity-related errors in the private-undetermined cell, while relatively low, was nonetheless greater than the number that occurred in a control-no-influence condition. More definitive evidence for the role of contact would require the comparison with a condition in which the group judgments were, for example, presented on videotape.

[p. 193 ↓]

Self-Relevant Individualistic and Collectivistic Values

Although contact and the twin concerns with being liked and being right may partially account for conformity effects, there is convincing evidence that these variables

do not provide a complete explanation. Bond and Smith (1996) conducted a meta-analysis of cross-cultural line-judgment studies similar to those used by Asch (1952, 1956) and found that conformity was related significantly to Hofstede's (1980) index of individualism-collectivism. Conformity was higher in collectivistic cultures.

What makes the Bond and Smith evidence particularly relevant to the present argument is the possibility that individualistic and collectivistic values are related to the self. Markus and Kitayama (1991) argue that members of individualistic cultures base self-esteem on the "ability to express self, validate internal attributes" (1991: 230, Table 1), while members of collectivistic cultures base self-esteem on the "ability to adjust, restrain self, maintain harmony with social context" (1991: 230, Table 1). Consistent with Markus and Kitayama's interpretation, Sedikides et al. (2003) report evidence relating the differential association of self-positivity to individualistic and collectivistic behaviors and traits. They compared the self-positivity of US college students of Japanese or non-Japanese backgrounds on reported hypothetical individualistic or collectivistic behaviors (e.g., "Put yourself before your group" or "Defend your group's decisions," [1991: 64, Table 1]) and traits (e.g., "independent" or "loyal," [1991: 64, Table 2]). Following Aliche's (1985) finding of a "better than average" effect, self-reports compared own behavior or traits with the behavior or traits of the supposed typical other group member. The results indicated that students of Japanese backgrounds saw themselves as relatively more positive than others to a greater extent on collectivistic than individualistic behaviors and traits while students of non-Japanese backgrounds saw themselves as relatively more positive than others to a greater extent on individualistic behaviors and traits than collectivistic behaviors and traits.

The Conformity Dilemma

Both the concern with being liked and the concern with being right could be interpreted as flowing from consistency with positive self-esteem and thus as implying that self-esteem should be positively correlated with conformity. In view of the persuasive evidence supporting sociometry theory, this possibility is particularly compelling to the extent that conformity flows from the concern with being liked.

Even though the evidence consistent with sociometry theory could be interpreted as implying that self-esteem should be positively associated with conformity, the literature on individualism and collectivism implies that for people with individualistic values self-esteem should be negatively associated with conformity. If both hypothesized processes occur, one would expect that for people with individualistic values the dilemma posed by conformity pressure would be particularly acute. The dilemma is between maintaining high self-esteem by giving in to conformity pressure and obtaining the assumed approval of the group or maintaining high self-esteem by resisting conformity pressure and behaving consistently with individualistic values. On the other hand, for individuals with collectivistic values, the dilemma posed by conformity pressures should be much less acute.⁴

There is indirect evidence that participants in Western conformity experiments do recognize the dilemma. Allen (1975: 18) reports some unpublished data indicating that perceived deviation from group consensus in the Johnny Rocco case (Schachter, 1951) led to the anticipation of being voted out of the group, and Sabini et al. (2001) cite Jahoda (1959) as reporting that both conforming and nonconforming participants in [p. 194 ↓] the Asch (1952, 1956) studies were upset by the experience.

Imbalance Resolution and Dissonance Reduction

Aronson (1968), Steele (1988), Steele and Liu (1983, and Schlenker (1982), have all seen the self as related to dissonance. However, none of these social psychologists extended their interpretations to include imbalance in self-related semicycles. The relevance of imbalance in self-related semi-cycles to dissonance can be illustrated with two examples from the literature on so-called free-choice situations – situations in which the participants chooses between two desirable alternatives.⁵

The first investigation of free-choice situations was a study by Brehm (1956). Consistent with the theoretical argument that a choice between desirable alternatives creates dissonance, Brehm reported that the chosen alternative increased in rated desirability

and the rejected alternatives decreased in rated desirability. Such so-called spreading of the alternatives was interpreted as evidence for dissonance reduction. However, later research by Shultz et al. (1999) found that the spreading was due, not to the chosen alternative increasing in desirability, but to the rejected alternative decreasing in desirability. Such a pattern is consistent with a balance interpretation.

Note that for a positive self (+) to choose (+) a desirable alternative (+) is consistent and thus does not provide a two-valued basis for change. However, for a positive self (+) to reject (-) a desirable alternative (+) is imbalanced and does provide a basis for change. The balance-achieving change could involve lowering self-esteem but that would create inconsistency in the self-concept. Thus the simplest route to imbalance resolution would involve decreasing the perceived desirability of the rejected alternative.⁶

Further evidence consistent with a balance interpretation of alternative spreading comes from Brock's (1963) finding that spreading of the alternatives was greater when the alternatives were objectively dissimilar than when the alternatives were objectively similar. Note quite simply that objective similarity is a positive unit relation and that when the alternatives are perceived as associated by a positive sign different evaluation of the alternatives would create an imbalanced semicycle.

Beyond Two-Valued to Many-Valued Relations – the Tetrahedron Model

Although logic is two-valued, human thought is obviously capable of many-valued distinctions. How then can logic's two-valued distinctions be extended to many-valued distinctions? One possibility is by combining logic and probability theory as in the probabilistic models of McGuire (1960) and Wyer (1974). A potential objection to these models is that they relied on traditional formulations of logic when, as argued above, it is possible that the multiplicative rule provides a better description of human thought than does the syllogism.

An approach that relies on a corollary of the multiplicative rule is Rosenberg's (1956, 1960), above-described technique of multiplying the bipolar ratings of the value and

instrumental cognition within a cognitive band. This approach was successful in demonstrating a correlation between the sum of such multiplications across cognitive bands and evaluation of the central attitude in the structure. However, insofar as the concern is with predicting a specific evaluation of the central attitude there is an obvious problem in that the multiplication will, in some instances, result in a product that goes off the scale. Fortunately, this problem does not occur with the approach of the tetrahedron model, a model first suggested by [p. 195 ↓] Wiest (1965) but given mathematical form by Wellens and Thistlethwaite (1971a, 1971b).

Consider each of the three relations in the *p-o-x* semicycle, not just as having either plus or minus values, but as having more than two values on a bipolar scale, for example -2, -1, 0, +1, +2. Wiest associated each of these relations with one of the left-right, down-up, and near-far dimensions of a cube. Any cube, of course, has eight corners, four on the "top" and four on the "bottom." Wiest assumed that the corners of the cube represented the pure, two-valued possibilities and thus that four of the corners (+ + +, + -, - + -, - +) were balanced and the remaining four were imbalanced. Connecting the balanced corners inside the cube creates a tetrahedron – a spatial form that can be thought of as a three-sided pyramid. Wiest generalized beyond the two-valued corners by assuming that any point on the surface of the tetrahedron or inside the tetrahedron was balanced, and thus suggested an elegantly simple solution to the perplexing problem as to how to generalize beyond two-valued possibilities.

It is, of course, arbitrary which relation is associated with which spatial dimension, but suppose that the *p-o* relation is associated with the down-up dimension with negative values down and positive values up. Connecting any two values of the *p-x* and *o-x* dimensions on the top of the cube creates a point that can be projected with a straight line into the cube. This line will intersect the tetrahedron at its upper boundary, proceed through the tetrahedron, and emerge from the tetrahedron at its lower boundary. Wellens and Thistlethwaite (1971a, 1971b) wrote one equation that, given any paired values for the *p-x* and *o-x* relations, predicts the upper-boundary intersection and a second equation that predicts the lower-boundary intersection.

Because Wiest assumed that any point on the surface of the tetrahedron or inside the tetrahedron was balanced, it follows that any predicted value for *p-o* from the upper boundary down to and including the lower boundary would be balanced. Since,

however, for paired p - x and o - x values that are not highly polarized the specified area inside the tetrahedron is large, the range of balanced possibilities is also large. A simple solution to this problem is to assume that the most probable predicted value is midway between the upper and lower boundaries, or is a simple average of the upper- and lower-boundary predictions. This approach to combining the upper- and lower-boundary predictions creates Wellens and Thistlethwaite's equal weights model.

Other models are, of course, possible, but Wellens and Thistlethwaite present only one, a so-called unequal weights model in which the upper-boundary prediction is weighted 0.75 and the lower-boundary prediction is weighted 0.25. How do the predictions for the equal and unequal weights models differ? If, as assumed above, the upper end of the p - o dimension represents the positive values, the predicted p - o values for less polarized p - x and o - x pairings are more positive or less negative for the unequal weights model than for the equal weights model. The unequal-weights model predicts less p to o disliking as a function of p and o disagreement than does the equal weights model.

Research by Wellens and Thistlethwaite (1971a, 1971b), and Tashakkori and Insko (1979, 1981) in which the p - o relation (attraction) was the dependent variable has supported the unequal weights model. Wellens and Thistlethwaite used a procedure in which participants played the role of a p who held each of five evaluations (-2, -1, 0, +1, +2) regarding a student demonstration, and considered an o who also held each of the five evaluations. Tashakkori and Insko used a variation of the Byrne (1971) anonymous stranger technique. In this latter research participants initially spent some time at a computer responding to a large number of attitude items that had been pilot tested to represent the different levels of the p - x variable. After random assignment of the participants to one cell of the p to x and o to x design the computer selected five x 's from each participant's prior responses, and asked the participants to rate their liking of the other person whose evaluations of these x 's were given. As indicated above, all of this research supported the unequal weights model.

Why with a p - o , or attraction, dependent variable should the unequal weights model have produced the superior fit? A possible explanation parallels the above postulated concern with being liked (or not being disliked). Recall that in the case of p and o disagreement the unequal weights model predicts less disliking than does the equal weights model. Note that p to o disliking should result in reciprocal o to p disliking

in the $p\text{-}o$, $o\text{-}p$ semicycle, and that transforming the o to p disliking relation into an element, being disliked, would produce imbalance in a self-relevant semicycle, I (+) am (+) disliked (-). Stated more simply, it is the reluctance to entertain the possibility of negative self-evaluation that results in the reluctance to assume extreme dislike of o.⁷

The problem with modeling the $p\text{-}o\text{-}x$ semicycle is familiar to examining a single semicycle in isolation from other semicycles that may be implicitly, or explicitly, involved. In research in which the p to o relation is the dependent variable the argument is that it is the reluctance to assume being disliked that shifts p to o sentiment toward the more positive, or less negative, upper boundary and thus provides a superior fit for the unequal weights model. The argument here parallels the above argument that the concern with being liked (or not disliked) partially accounts for conformity effects and for agreement effects in hypothetical semicycles.

Tashakkori and Insko (1981) reasoned that they could block the tendency to avoid implied dislike by manipulating o to p and testing for o to x . A single experiment found, as expected, that the results fit the equal weights model.

Although these results are preliminary and research with other semicycles should be conducted, the findings do suggest that the tetrahedron model has promise. Understanding how to extend the multiplicative rule to other than two-valued distinctions and understanding how to conceptualize low self-regard are the two outstanding problems confronting further theoretical development.

Notes

1 Philosophers use the term “fuzzy logic” to refer to a controversial form of many-valued logic (cf. Haack, 1978: 165–167; Haack, 1996: 229–258). Haack (1978) characterizes fuzzy logic as partially related to fuzzy set theory according to which membership in a set is not restricted to present or absent but a matter of degree represented by a real number between 0 and 1. For example, if person A belongs to the degree 0.2 to the set of tall people, it follows in fuzzy logic that the proposition “A is tall” has the value 0.2, or has a low degree of truth. Haack (1996: 230) stated that ‘I remain convinced,

first ... that truth does not come in degrees, and, second that fuzzy logic is not a viable competitor of classical logic."

2 Despite the fact that low self-esteem is only low in a relative sense, and may be more characteristic of a short-term "state" than a long-term "trait" (Heatherton and Polivy, 1984), low self-esteem is nonetheless, as Baumeister (1995) suggests, a "puzzle." This puzzle is amply illustrated by Swann et al.'s (1987) finding that while participants low in social self-esteem consistently judged unfavorable feedback regarding their body language as more valid than favorable feedback, they reacted to such unfavorable feedback with relatively more negative affect. Such inconsistency between affect and cognition is indeed a puzzle and merits further study. Does the negative affect relate to the implied difficulty of bolstering long-term, trait self-esteem?

3 In view of the fact that responses to hypothetical situations may differ from responses to nonhypothetical situations, some readers may wonder why suggestions regarding hypothetical *p-o-x* semicycles merit investigation. This question has at least two answers. First, in view of the fact that the influential Zajonc critique (1968) was based on results from the study of hypothetical *p-o-x* semicycles, it is theoretically important to thoroughly explore those semicycles. Second, although nonhypothetical situations may be of more practical significance than hypothetical situations, balance theory is a theory regarding thought processes regardless as to whether those processes relate to practical-nonhypothetical situations or impractical-hypothetical situations.

[p. 197 ↓] 4 Since everyone has learned to rely on the evidence of their senses, the dilemma, even though less acute for individuals with collectivistic values, should still be present.

5 For evidence relating to the insufficient-reward effect (Festinger and Carlsmith, 1959) see Insko et al. (1975).

6 The inconsistency with positive self-interpretation has an interesting fit with Baumeister et al.'s (2001: 323) extensive documentation that "bad is stronger than good" and that self-protective motivation is stronger than self-enhancement motivation.

7 The above interpretation of the superior fit of the unequal weights model does require that a relation be transformed into an element, but since ordinary speech routinely transforms verbs into nouns this requirement does not pose any particular difficulty. Note, further, that the interpretation does not require that the participant go through an explicit reasoning process. Although the reasoning process could be explicit, it is more likely that reasoning regarding any self-relevant semicycle is implicit and automatic. The argument parallels Aristotle's (1945) above described assumption that in deducing the implications of a practical syllogism we do not bother with an explicit consideration of self-relevant premises.

References

- Abelson, R.P. (1959) Modes of resolution of belief dilemmas . Journal of Conflict Resolution , 3, 343–352. <http://dx.doi.org/10.1177/002200275900300403>
- Abelson, R.P. (1983) Whatever became of consistency theory . Personality and Social Psychology Bulletin , 9, 37–54. <http://dx.doi.org/10.1177/0146167283091006>
- Abelson, R.P. and Rosenberg, M.J. (1958) Symbolic psycho-logic: A model of attitudinal cognition . Behavioral Science , 62, 990–1008.
- Aderman, D. (1969) Effects of anticipating future interaction on the preference for balanced states . Journal of Personality and Social Psychology , 11, 214–219. <http://dx.doi.org/10.1037/h0027028>
- Allen, V.L. (1975) Social support for nonconformity . In L. Berkowitz (ed.) Advances in Experimental Social Psychology , 8, 1–43. New York: Academic Press.
- Alicke, M.D. (1985) Global self-evaluation as determined by the desirability and controllability of trait adjectives . Journal of Personality and Social Psychology , 49, 1621–1630. <http://dx.doi.org/10.1037/0022-3514.49.6.1621>
- Anderson, N.H. (1981) Integration theory applied to cognitive responses and attitudes . In R.E. Petty, ed. , T.M. Ostrom, ed. and T.C. Brock (eds), Cognitive Responses in Persuasion . Hillsdale, NJ: Erlbaum.

Aristotle (1945) Movement of Animals (E.S. Forster, ed. , Trans.). Cambridge: Harvard University Press.

Aronson, E. (1968) Dissonance theory: Progress and problems . In R.P. Abelson, ed. , E. Aronson, ed. , W.J. McGuire, ed. , T.M. Newcomb, ed. , M.J. Rosenberg, ed. and P.H. Tannenbaum (eds), Theories of Cognitive Consistency: A Sourcebook , pp. 5–27. Chicago: Rand McNally.

Aronson, E. and Cope, V. (1968) My enemy's enemy is my friend . Journal of Personality and Social Psychology , 8, 8–12. <http://dx.doi.org/10.1037/h0021234>

Aronson, E. and Worchel, P. (1966) Similarity versus liking as determinants of interpersonal attractiveness . Psychonomic Science , 5, 157–158.

Asch, S.E. (1952) Social Psychology . Englewood Cliffs, NJ: Prentice-Hall. <http://dx.doi.org/10.1037/10025-000>

Asch, S.E. (1956) Studies of independence and conformity: A minority of one against a unanimous majority . Psychological Monographs , 70(46).

Baumeister, R.F. (1982) A self-presentational view of social phenomena . Psychological Bulletin , 91, 3–26. <http://dx.doi.org/10.1037/0033-2909.91.1.3>

Baumeister, R.F. (2011) Self Esteem: The Puzzle of Low Self-regard . New York: Plenum Press.

Baumeister, R.F., Bratslavsky, E., Finkenauer, C. and Vohs, K.D. (2001) Bad is stronger than good . Review of General Psychology , 5, 323–370. <http://dx.doi.org/10.1037/1089-2680.5.4.323>

Baumeister, R.F., Tice, D.M. and Hutton, D.G. (1989) Self-presentation motivations and personality differences in self-esteem . Journal of Personality , 56, 571–598. <http://dx.doi.org/10.1111/j.1467-6494.1988.tb00903.x>

Benesh-Weiner, M. (ed.) (1988) Fritz Heiden The Notebooks: Vol. 4: Balance Theory . Weinheim: Springer-Verlag.

Berscheid, E. and Walster, P. (1978) *Interpersonal Attraction* . Menlo Park, CA: Addison-Wesley.

Bond, R. and Smith, P.M. (1996) Culture and conformity: A meta-analysis using Asch's (1952b, 1956) line judgment task . *Psychological Bulletin* , 119, 111–137. <http://dx.doi.org/10.1037/0033-2909.119.1.111>

Brehm, J. (1956) Postdecision changes in the desirability of alternatives . *Journal of Abnormal and Social Psychology* , 60, 379–383.

Brock, T.C. (1963) Effects of prior dishonesty on post-decision dissonance . *Journal of Abnormal and Social Psychology* , 66, 325–331. <http://dx.doi.org/10.1037/h0047653>

Brock, T.C. (1967) Communication discrepancy and intent to persuade as determinants of counterargument production . *Journal of Experimental Social Psychology* , 3, 296–309. <http://dx.doi.org/10.1016/0022-1031%2867%2990031-5>

Bruner, G., Goodnow, J.J. and Austin, G. (1956) *A Study of Thinking* . New York: Wiley.

Byrne, D. (1971) *The Attraction Paradigm* . New York: Academic Press.

Byrne, D. and Griffitt, W. (1966) Similarity versus liking: A clarification . *Psychonomic Science* , 6, 295–296.

Byrne, D. and Rhamey, R. (1965) Magnitude of positive and negative reinforcements as a determinant of attraction . *Journal of Personality and Social Psychology* , 2, 884–889. <http://dx.doi.org/10.1037/h0022656>

Carnegie, D. (1937) *How to Win Friends and Influence People* . New York: Simon and Schuster.

Cartwright, D. and Harary, F. (1956) Structural balance: A generalization of Heider's theory . *Psychological Review* , 63, 277–293. <http://dx.doi.org/10.1037/h0046049>

Cialdini, R.B. (2009) *Influence: Science and Practice* . New York: Allyn and Bacon

Cohen, M.R. (1944) *A Preface to Logic* . New York: Holt.

Deutsch, M. and Gerard, H.B. (1955) A study of normative and informational social influences upon individual judgment . *Journal of Abnormal and Social Psychology* , 51, 629–636. <http://dx.doi.org/10.1037/h0046408>

Festinger, L. (1950) Informal social communication . *Psychological Review* , 57, 271–282. <http://dx.doi.org/10.1037/h0056932>

Festinger, L. (1954a) A theory of social comparison processes . *Human Relations* , 7, 117–140. <http://dx.doi.org/10.1177/001872675400700202>

Festinger, L. (1954b) Motivation leading to social behavior . In M.R. Jones (ed.), *Nebraska Symposium on Motivation* , 2, 121–218. Lincoln, NE: University of Nebraska Press.

Festinger, L. and Carlsmith, J.M. (1959) Cognitive consequences of forced compliance . *Journal of Abnormal and Social Psychology* , 58, 203–211. <http://dx.doi.org/10.1037/h0041593>

Fishbein, M. (1963) An investigation of the relationships between beliefs about an object and the attitude toward that object . *Human Relations* , 16, 233–240. <http://dx.doi.org/10.1177/001872676301600302>

Gawronski, B. and Walther, E. (2008) The TAR effect: When the ones who dislike become the ones who are disliked . *Personality and Social Psychology Bulletin* , 34, 1276–1289. <http://dx.doi.org/10.1177/0146167208318952>

Gilovich, T. (1991) *How We Know What Isn't So* . New York: Free Press.

Greenwald, A.G. (1967) An amended learning model of persuasion . Paper presented at the meeting of the American Psychological Association, Washington, DC.

Greenwald, A.G. (1968) Cognitive learning, cognitive responses to persuasion, and attitude change . In A.G. Greenwald, ed. , T.C. Brock, ed. and A.T.M. Ostrom (eds), *Psychological Foundations of Attitudes* . New York: Academic Press.

Haack, S. (1978) *Philosophy of Logics* . Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511812866>

Haack, S. (1996) *Deviant Logic, Fuzzy Logic: Beyond the Formalism* . Chicago: University of Chicago Press.

Harary, R. (1983) Consistency theory is alive and well . *Personality and Social Psychology Bulletin* , 9, 60–64. <http://dx.doi.org/10.1177/0146167283091008>

Heider, F. (1946) Attitudes and cognitive organization . *Journal of Psychology* , 21, 107–112. <http://dx.doi.org/10.1080/00223980.1946.9917275>

Heider, F. (1958) *The Psychology of Interpersonal Relations* . New York: Wiley. <http://dx.doi.org/10.1037/10628-000>

Henle, M. (1962). On the relation between logic and thinking . *Psychological Review* , 9, 366–378. <http://dx.doi.org/10.1037/h0042043>

Hofstede, G. (1980) *Cultural Consequences: International Differences in Work-related Values* . Beverly Hills, CA: Sage.

Insko, C.A. (1981) Balance theory and phenomenology . In R. Petty, ed. , T. Ostrom, ed. and T. Brock (eds), *Cognitive Responses in Persuasion* , pp. 309–338. New Jersey: Erlbaum.

Insko, C.A. (1984) Balance theory, the Jordan paradigm, and the Wiest tetrahedron . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 18, 89–140. New York: Academic Press.

Insko, C.A. (1999) A balance-logic perspective on Kruglanski and Thompson's single-route approach to persuasion . *Psychological Inquiry* , 10, 127–137. <http://dx.doi.org/10.1207/S15327965PL100206>

Insko, C.A., Blake, R.R., Cialdini, R.B. and Muliak, S.A. (1970) Attitude toward birth control and cognitive consistency: Theoretical and practical implications of survey data .

Journal of Personality and Social Psychology , 16, 228–237. <http://dx.doi.org/10.1037/h0029826>

Insko, C.A., Drenan, S., Solomon, M.R., Smith, R.H. and Wade, T. (1983) Conformity as a function of the consistency of positive self-evaluation with being liked and being right . Journal of Experimental Social Psychology , 19, 341–358. <http://dx.doi.org/10.1016/0022-1031%2883%2990027-6>

Insko, C.A., Smith, R.H., Alicke, M.D., Wade, J. and Taylor, S. (1985) Conformity and group size: The concern with being right and the concern with being liked . Personality and Social Psychology Bulletin , 11, 41–50. <http://dx.doi.org/10.1177/0146167285111004>

Insko, C.A., Thompson, V.D., Wtroebe, W., Shaud, K.F., Pinner, B.E. and Layton, B.D. (1973) Implied evaluation and the similarity-attraction effect . Journal of Personality and Social Psychology , 25, 297–308. <http://dx.doi.org/10.1037/h0034224>

Insko, C.A. and Wilson, M. (1977) Interpersonal attraction as a function of social interaction . Journal of Personality and Social Psychology , 35, 903–911. <http://dx.doi.org/10.1037/0022-3514.35.12.903>

Insko, C.A., Worchel, S., Folger, R. and Kutkus, A. (1975) A balance theory interpretation of dissonance . Psychological Review , 82, 169–183. <http://dx.doi.org/10.1037/0033-295X.82.3.169>

James, W. (1890) Principles of Psychology . New York: Dover. <http://dx.doi.org/10.1037/11059-000>

Jahoda, M. (1959) Conformity and obedience . Human Relations , 12, 99–120. <http://dx.doi.org/10.1177/001872675901200201>

Jordan, N. (1953) Behavioral forces that are a function of attitude and of cognitive organization . Human Relations , 6, 273–287. <http://dx.doi.org/10.1177/001872675300600304>

- LaPrelle, J., Hoyle, R.H., Insko, C.A. and Bernthal, P. (1990) Interpersonal attraction and descriptions of the traits of others: Ideal similarity, self similarity, and liking . *Journal of Research in Personality* , 24, 216–240. <http://dx.doi.org/10.1016/0092-6566%2890%2990018-2>
- LaPrelle, J., Insko, C.A., Cooksey, L. and Graetz, K. (1991) Self similarity, ideal similarity, and undesired self similarity as indicators of attraction . *Representative Research in Personality* , 19, 65–74.
- Leary, M.R. and Baumeister, R.F. (2000) The nature and function of self-esteem: sociometer theory . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 32, 1–62. New York: Academic Press.
- Leary, M.R., Gallagher, B., Fors, E., Buttermore, N., Baldwin, E., Kennedy, K. et al. (2003) The invalidity of disclaimers about the effects of social feedback on self-esteem . *Personality and Social Psychology Bulletin* , 5, 623–636. <http://dx.doi.org/10.1177/0146167203029005007>
- Leary, M.R. and Tangney, J.P. (2003) The self as an organizing construct in the behavioral and social sciences . In M.R. Leary, ed. and J.P. Tangney (eds), *Handbook of Self and Identity* , pp. 3–14. New York: Guilford.
- Markus, H.R. and Kitayama, S. (1991) Culture and the self: Implications for cognition, emotion and motivation . *Psychological Review* , 98, 224–253. <http://dx.doi.org/10.1037/0033-295X.98.2.224>
- McGuire, M.J. (1960) A syllogistic analysis of cognitive relationships . In C.I. Hovland, ed. and M.J. Rosenberg (eds), *Attitude Organization and Change: An Analysis of Consistency Among Attitude Components* , pp. 65–111. New Haven, CT: Yale University Press.
- Miller, N.E. (1951) Learnable drives and rewards . In S.S. Stevens (ed.), *Handbook of Experimental Psychology* pp. 435–472. New York: Wiley.
- Mill, J.S. (1874) *A System of Logic* . New York: Harper.

Mita, T.H., Dermer, M. and Knight, J. (1977) Reversed facial images and the mere exposure hypothesis . *Journal of Personality and Social Psychology* , 35, 597–601. <http://dx.doi.org/10.1037/0022-3514.35.8.597>

Montoya, M. and Insko, C.A. (2007) Toward a more complete understanding of the reciprocity of liking effect . *European Journal of Social Psychology* , 38, 477–498. <http://dx.doi.org/10.1002/ejsp.431>

Newcomb, T.M. (1968) Interpersonal balance . In R.P. Abelson, ed. , E. Aronson, ed. , W.J. McGuire, ed. , T.M. Newcomb, ed. , M.J. Rosenberg, ed. and P.H. Tannenbaum (eds), *Theories of Cognitive Consistency: A Sourcebook* . Chicago: Rand McNally.

Osgood, C.E. and Tannenbaum, P.H. (1955) The principle of congruity in the prediction of attitude change . *Psychological Review* , 62, 42–55. <http://dx.doi.org/10.1037/h0048153>

Peak, R. (1955) Attitude and motivation . In M.R. Jones (ed.), *Nebraska Symposium on Motivation* , 3, 149–188. Lincoln: University of Nebraska Press.

Rosenberg, M. (1979) *Conceiving the Self* . New York: Basic Books.

Rosenberg, M.J. (1956) Cognitive structure and attitudinal affect . *Journal of Abnormal and Social Psychology* , 53, 367–372. <http://dx.doi.org/10.1037/h0044579>

Rosenberg, M.J. (1960) An analysis of affective-cognitive consistency . In C.I. Hovland (ed.), *Attitude Organization and Change* , 15–64. New Haven: Yale University Press.

Runkel, P.J. and Peizer, D.B. (1968) The two-valued orientation of current equilibrium theory . *Behavioral Science* , 13, 56–65. <http://dx.doi.org/10.1002/bs.3830130107>

Saegert, S., Swap, W. and Zajonc, R.B. (1973) Exposure, context, and interpersonal attraction . *Journal of Personality and Social Psychology* , 25, 234–242. <http://dx.doi.org/10.1037/h0033965>

- Sabini, J., Siepmann, M. and Stein, J. (2001) The really fundamental attribution error in social psychological research . *Psychological Inquiry* , 12, 1–15. http://dx.doi.org/10.1207/S15327965PLI1201_01
- Schachter, S. (1951) Deviation, rejection, and communication . *Journal of Abnormal and Social Psychology* , 46, 190–207. <http://dx.doi.org/10.1037/h0062326>
- Schlenker, B.R. (1982) Translating actions into attitudes: An identity-analytic approach to the explanation of social conduct . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 15, 193–247. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960298-1>
- Schmitt, D.P. and Allik, J. (2005) Simultaneous administration of the Rosenberg self-esteem scale in 53 nations: Exploring the universal and culture-specific features of global self-esteem . *Journal of Personality and Social Psychology* , 89, 623–642. <http://dx.doi.org/10.1037/0022-3514.89.4.623>
- Sedikides, C. (1993) Assessment, enhancement, and verification determinations of the self-evaluation process . *Journal of Personality and Social Psychology* , 65, 317–338. <http://dx.doi.org/10.1037/0022-3514.65.2.317>
- Sedikides, C., Gaertner, L. and Toguchi, Y. (2003) Pancultural self-enhancement . *Journal of Personality and Social Psychology* , 84, 60–79. <http://dx.doi.org/10.1037/0022-3514.84.1.60>
- Sedikides, C. and Skowronski, J.J. (1997) The symbolic self in evolutionary development . *Personality and Social Psychology Review* , 1, 80–102. http://dx.doi.org/10.1207/s15327957pspr0101_6
- Srivastava, S. and Beer, J. (2005) How self-evaluations relate to being liked by others: Integrating sociometer and attachment perspectives . *Journal of Personality and Social Psychology* , 89, 966–977. <http://dx.doi.org/10.1037/0022-3514.89.6.966>
- Shultz, T.R., Leveille, E. and Lepper, M.R. (1999) Free choice and cognitive dissonance revisited: Choosing ‘lesser evils’ versus ‘greater good’ . *Personality and Social Psychology Bulletin* , 25, 40–48. <http://dx.doi.org/10.1177/0146167299025001004>

Steele, C.M. (1988) The psychology of self-affirmation: Sustaining the integrity of the self . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 21, 261–302. San Diego: Academic Press.

Steele, C.M. and Liu, T.J. (1983) Dissonance processes as self-affirmation . *Journal of Personality and Social Psychology* , 45, 5–19. <http://dx.doi.org/10.1037/0022-3514.45.1.5>

Stroebe, W., Thompson, V.D., Insko, C.A. and Reisman, S.R. (1970) Balance and differentiation in the evaluation of linked attitude objects . *Journal of Personality and Social Psychology* , 16, 48–54. <http://dx.doi.org/10.1037/h0030062>

Swann Jr., W.B., Griffin, J.J., Predmore, S.C and Gaines, B. (1987) The cognitive-affective crossfire: When self-consistency confronts self-enhancement . *Journal of Personality and Social Psychology* , 52, 881–889. <http://dx.doi.org/10.1037/0022-3514.52.5.881>

Tashakkori, A. and Insko, C.A. (1979) Interpersonal attraction and the polarity of similar attitudes: A test of three balance models . *Journal of Personality and Social Psychology* , 37, 2262–2277. <http://dx.doi.org/10.1037/0022-3514.37.12.2262>

Tashakkori, A. and Insko, C.A. (1981) Interpersonal attraction and person perception: Two tests of three balance models . *Journal of Experimental Social Psychology* , 17, 266–285. <http://dx.doi.org/10.1016/0022-1031%2881%2990028-7>

Taylor, S.E. and Brown, J.D. (1988) Illusion and well-being: A social psychological perspective on mental health . *Psychological Bulletin* , 103, 193–210. <http://dx.doi.org/10.1037/0033-2909.103.2.193>

Wellens, A.R. and Thistlethwaite, D.L. (1971a) An analysis of two quantitative theories of cognitive balance . *Psychological Review* , 78, 141–150. <http://dx.doi.org/10.1037/h0030666>

Wellens, A.R. and Thistlethwaite, D.L. (1971b) Comparison of three theories of cognitive balance . *Journal of Personality and Social Psychology* , 20, 89–92. <http://dx.doi.org/10.1037/h0031704>

Wertheimer, M. (1923) Untersuchungen zur Lehre von der Gestalt . Psychologische Forschung , 4, 301–350. <http://dx.doi.org/10.1007/BF00410640>

Wetzel, C. and Insko, C.A. (1974) Preacquaintance attraction as an interactive function of the proportion and number of similar attitudes . Representative Research in Social Psychology , 1974, 27–34.

Wiest, W.M. (1965) A quantitative extension of Heider's theory of cognitive balance applied to interpersonal perception and self-esteem . Psychological Monographs , 79 (14, Whole No. 607).

Woodruff, H. and DiVesta, F. (1948) The relationship between values, concepts, and attitudes . Educational and Psychological Measurement , 8, 645–660. <http://dx.doi.org/10.1177/001316444800800409>

Wyer, R.S. (1974) Cognitive Organization and Change: An Information Processing Approach . Hillsdale, NJ: Lawrence Erlbaum Associates.

Zajonc, R.B. (1968) Cognitive theories in social psychology . In G. Lindzey, ed. and E. Aronson (eds), Handbook of Social Psychology , 1, 320–411. Reading, MA: Addison-Wesley.

Zajonc, R.B. (2001) Mere exposure: A gateway to the subliminal . Current Directions in Psychological Science , 10, 224–227. <http://dx.doi.org/10.1111/1467-8721.00154>

<http://dx.doi.org/10.4135/9781446249215.n10>

Handbook of Theories of Social Psychology: Volume 1

Lay Epistemic Theory

Contributors: Arie W. Kruglanski

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Lay Epistemic Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n11>

Print pages: 201-224

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n11>

[p. 201 ↓]

Chapter 10: Lay Epistemic Theory

Arie W. Kruglanski, ed.

Abstract

This chapter traces the development of the theory of lay epistemics as it has unfolded over the last three decades. The chapter describes the serendipitous beginnings of the theory, its roots in twentieth-century philosophy and sociology of science, and its reaction to social cognitive theory on attribution and consistency processes. The general theoretical depiction of the knowledge formation process gave rise to three streams of lay epistemic research described herein. The respective research programs revolve around (1) the evidential base of knowledge formation giving rise to research on the *unimodel*, (2) the unique sociopsychological nature of evidence giving rise to research on *epistemic authority*, and (3) the motivational basis of knowledge formation giving rise to research on the *need for cognitive closure*. The evolution of these research programs is described in detail as are the factors affecting the theoretical decisions that lend direction to the lay epistemic framework. Because subjective knowledge forms the basis of a wide variety of social behavior the lay epistemic theory is relevant to a considerable range of real world concerns, a sample of which is described here.

Introduction

Thinking back to the origins of my lay epistemic theory, I can discern three major sources of influence on its formation and development. One source was my PhD

advisor Harold Kelley who, in a passing chat with me in 1968, mentioned an interest in investigating something he labeled as “lay epistemology.” I didn’t think about it much at the time but the term stuck in my mind, apparently. This constituted the first source of influence. Years later, as a senior lecturer at Tel Aviv University, I became strongly interested in philosophy of science and issues of the scientific method (owing partly to the influence of another mentor, Barry Collins, a UCLA professor and a devoted disciple of Donald Campbell). Leaving for a sabbatical at the University of North Carolina, I took along a bunch of philosophy of science books and was reading them voraciously in my free moments. I was particularly struck by the ideas of Karl Popper and his notion [p. 202 ↓] that the construction of scientific knowledge is similar in essence to the construction of lay knowledge, and hence that science is “common sense writ large” as he succinctly put it (Popper, 1959: 22). That notion resonated with my latent interest in pursuing the study of lay epistemology, stimulated by Kelley’s casual remark,¹ and I was particularly taken by Popper’s portrayal of the scientific process as a possible model of the way lay persons go about constructing their knowledge.

The third source of influence stemmed from a conviction, also indebted to Popper’s impact, that an aspirational ideal in theory construction is the maxim of generality. In Popper’s terms, scientific theories ought to be bold and comprehensive; that is, as rich as possible in empirical contents creating a correspondingly large set of opportunities for falsification. This precept that I found compelling² evoked my concern that the work in which I was engaged at the time just wasn’t general enough and that it ultimately targeted surface structures rather than deep structures, eluding the underlying principles that govern manifest phenomena. An explanation of how this transpired requires a slight detour in which that line of work is briefly described.

Internal-External or Endogenous-Exogenous Attributions?

My earlier line of work had to do with my critique of the internal-external distinction in attribution theory and my proposal to replace it with the endogenous-exogenous distinction (Kruglanski, 1975). Whereas the internal-external distinction juxtaposed

the *person* and the *situation* as the essential attributional categories, the endogenous-exogenous distinction juxtaposed actions performed for their own sake, constituting *ends in themselves*, and actions performed for ulterior ends, serving as *means* to further goals. I argued that both the interest in the task and desire for the external reward are equally *internal* to the actor, whereas both the task as such and the external reward are equally *external*. It made more sense to me that when participants perform a task for its interest value and appeal, they treat the task as an end in itself. I called their reasons in such case *endogenous* to the task. Conversely, when participants perform the task because they expect a reward or fear a punishment they treat the task as a means to an end. I called their reasons in such case as *exogenous* to the task.

The endogenous-exogenous distinction explained considerable amount of attributional findings and offered numerous possibilities for further inquiry, described in a comprehensive *Psychological Review* article (Kruglanski, 1975) and related empirical papers (Kruglanski et al., 1972, 1975a, 1975b). Indeed, for a while it seemed that the endogenous-exogenous theory would become the launching pad for my research program for years to come, and that I would continue exploring its implications and highlighting its insights into a wide range of social behaviors. This wasn't meant to pass, however, for soon after the publication of the endogenous-exogenous paper, I became troubled by the following thought: that besides the endogenous-exogenous and internal-external attributional distinctions, one can propose further and further distinctions that could possibly mediate human behavior in various domains. Thus, Weiner (1979, 1985) distinguished between attributions of performance outcomes to "effort," "ability," "luck," and "task difficulty," and creatively explored the implications of these attributional categories for achievement behavior. But the "scary" thought for me was that one could actually make attributions to an open-ended host of potential categories having to do with individuals' traits (e.g., aggressiveness, intelligence, honesty, friendliness), motivational and emotional states (e.g., fear, fatigue, greed, love, anxiety), environmental conditions (e.g., temperature, crowding, availability of resources), gender, ethnicity, and so on. All these might serve as [p. 203 ↓] useful attributional categories for some people in some circumstances, but somehow their exploration didn't appear to contribute much to the illumination of the ways in which people make causal assignments – which is what I thought attribution theory was all about!

This realization led me to the distinction between the *process* of causal attribution and the attributional *contents*. The two were insufficiently distinguished in prior attributional writings (e.g., Kelley, 1967), probably because most attributional analyses centered around a limited number of content categories (e.g., between internal and external attributions) that meshed easily with process elements. But once it was considered that attributional contents are innumerable, whereas the attributional process is unitary by definition (constituting the processes whereby causality generally is ascribed), the process-content distinction came into high relief (for me, that is). Because of a commitment to general principles, I quickly centered my efforts on delineating the “pure” attribution *process* stripped from the attributional *contents* attached to it in former publications.

Presumably such process was contained in the “attributional criteria” described in Kelley’s (1967) famous “ANOVA cube” consisting of the “consensus,” “consistency,” and “distinctiveness” rules. Specifically, according to Kelley, an attribution to an “entity” was assumed to be made if the effect (to be attributed) was distinctive to the entity and absent for other entities (the distinctiveness criterion), was consistently present across times and modalities when the entity was present (the consistency criterion), and was similarly produced by all relevant actors (the consensus criterion). But wait just a moment! Do these criteria depict a pure, content-free, attribution *process*? A moment’s thought reveals that they do not. Instead, they pertain to a situation wherein the attributor is deciding among specific potential content categories of causality: the *entity*, the *person*, the *time*, and the *modality*. Thus, an entity attribution is warranted where the effect *covaries* with the entity, and it *does not covary* with time, modality and person. Assuming that these four content categories need not be of universal interest to individuals and could be readily replaced by alternative causal categories about which one might be concerned, what is left then as an invariant process element is the principle of covariation. Indeed, in a subsequent publication Kelley (1971) identified this particular principle as the main vehicle of causal assignment, abandoning (so it would seem) the attributional criteria central to his prior exposition.

But why should covariation be evidence for causality? Very simply, it is part of the common definition of what a cause is. A cause (in common understanding) is something that covaries with the effect. Put differently, most people probably subscribe to the belief that “*if* something is a cause of an effect *then* it covaries with it.” This premise affords

the derivation that “if something does not covary with the effect it is NOT its cause.” It is precisely that kind of reasoning that allowed Kelley (1967) to rule out the *time*, *modality*, and *person* categories as causes of the effect, assigning the causal status to the only candidate category that exhibited covariation.

The astute reader would notice that the foregoing is an example of syllogistic reasoning. One starts with a major premise (concerning in this case covariation, a concept derived from the construct of causality), combines it with a minor premise (say, that a given entity exhibited covariation with some effect of interest) to deduce a conclusion, that the entity is (probable) cause of the effect. This framing immediately begged the question whether the covariation principle, apparently a major vehicle in the rendition of causal inferences isn't a special case of a broader principle involved in all possible inferences.

The syllogistic structure of reasoning (underlying the deduction of causality from covariation evidence) recalled Popper's [p. 204 ↓] assertion that all scientific knowledge is arrived deductively. Combined with Popper's notion that lay knowledge is arrived in the same manner as scientific knowledge, this notion of syllogistic reasoning, or what psychological researchers (e.g., Sloman, 1996) referred to as “rule following,” represents the essential principle whereby humans arrive at all their knowledge. This realization posed the challenge of transcending the realm of attribution processes, and to refocusing my research program on the broader issue of knowledge construction, marking the start of my theorizing about lay epistemics.³

Specifically, my analysis of causal inference was generalized to a major assumption of lay epistemic theory: that all knowledge⁴ is positively derived from evidence. In this respect, my reasoning differed from Popper's who insisted that knowledge can only be *falsified* by evidence. Let me explain. Popper (1959) assumed that the process of hypothesis testing is represented by the premise: “If H (hypothesis) is true then E (evidence) should be the case,” which implies that one can only disconfirm a hypothesis via a logical *modus tollens* (if E is false, H too must be false), but not verify it by evidence as I am suggesting. But, it is plausible that the knower may depart from a different (and stronger) assumption; specifically, the assumption that *if and only if* hypothesis H were true evidence E would be observed. The *if and only if* framing implies a bidirectional implication (in the terminology of logic known as the *equivalence*

relation) asserting that *both* “if H then E” and “if E then H” are true. For instance, one may feel confident (to the point of trusting it with one's life!) that the road is clear if the evidence of her or his eyes indicated this. The knower's premise in this instance may be expressed as “if and only if my eyesight indicates that the road is clear then it is indeed clear”.

In other words, given the *if and only if* premise (that may “feel” justifiable under specific circumstances) an individual could logically deduce the hypothesis from the evidence in a *modus ponens* fashion, whereby if E is affirmed (the evidence is observed) then H is inferred (the hypothesis is supported). Of course, the *if and only if* assumption might need to be modified on the basis of subsequent information which would cast doubt on the original conclusion that H was proven to be valid. For instance, if a rival alternative hypothesis H

1

appeared plausible and the need arose to distinguish it from the original H, one would have to formulate a subsequent inference rule whereby *if and only if H but not H*

1

were true then E

1

would obtain, and so on.⁵

Terminating the Epistemic Sequence

Philosophers of knowledge have long maintained that the sequence of hypothesis generation and testing (whether in science or lay belief formation) has no unique, or objective point of termination. In principle, it could continue interminably as one keeps engendering further and further alternative hypotheses without ever crystallizing firm knowledge on a topic. Obviously, such epistemic “obsession” would be highly dysfunctional and paralyzing. Most of the people, most of the time are quite capable of forming judgments based on available evidence, and of self-regulating adaptively on the basis of those judgments.

The theory of lay epistemics identifies two mechanisms that effect the cessation of the hypothesis generation sequence and the crystallization of confident knowledge; one of these mechanisms is *motivational*, the other is *cognitive*. The motivational mechanism is based on the need for cognitive closure (Kruglanski, 2004). Two types of such need have been distinguished, referred to respectively as the needs for *nonspecific* and *specific* closure. The need for nonspecific closure denotes a desire for a firm answer to a question, any firm answer as compared to confusion and ambiguity. The need for a specific closure denotes a specific desirable answer to a question; for example, an esteem [p. 205 ↓] enhancing answer, or an optimistic answer. Each of these needs is assumed to vary in degree and lie on a continuum ranging from a low to a high motivational magnitude. Thus, one may desire nonspecific closure strongly, mildly or not at all, actually craving to avoid it. Similarly, one may desire to reach a particular conclusion (or specific closure) with varying degrees of strength. Both types of need determine the length of the epistemic sequence of hypothesis generation and testing. The higher the need for nonspecific closure the shorter the sequence and the stronger the tendency to “seize and freeze” on accessible, closure-affording evidence. The higher the need for a specific closure, the stronger the tendency to terminate the sequence when the available evidence appears to yield the desired conclusion, or to keep the sequence going until such conclusion seems implied by the evidence (Ditto and Lopez, 1992).

The cognitive mechanism relates to the “authority” of the evidence, or what we have referred to as *epistemic authority*. (Kruglanski et al., 2005). Specifically, the theory assumes that some kinds of evidence are “incontrovertible,” because their source is deemed indubitable and beyond reproach. For instance, one might trust the evidence yielded by one's senses (e.g., seeing something with one's own eyes), or evidence provided by a trusted expert (one's physician, an auto mechanic, a parent, an admired leader priest or cleric), or evidence derived from “common knowledge,” that is from one's group's shared reality (Hardin and Higgins, 1996).

Explorations in Lay Epistemics

The transition of my research interests from attribution processes to those of knowledge formation was described in a *Psychological Review* article (Kruglanski, 1980) that

followed (by five years) the endogenous-exogenous theory described earlier. What ensued was three decades of research on various facets of lay epistemics organized around three major research programs. These addressed, respectively: (1) the *unimodel* of social judgment, (2) the construct of *epistemic authority*, and (3) the *need for cognitive closure*. Let me now describe the major theoretical and empirical developments in each of those domains.

The Unimodel of Human Judgment: All You Need is One! Lay Epistemic Theory and the Dual Process Models

As already noted, the lay epistemic theory assumes that all knowledge derives from evidence, broadly conceived. To reiterate, in constructing new knowledge, or in forming a new judgment, the individual is assumed to use an inference rule whereby if a given evidence E obtains, conclusion C follows (or hypothesis H is supported).

The (lay epistemic) notion that all knowledge is derived via unitary, syllogistic, process seemed inconsistent with the popular claim that judgments can be formed via two qualitatively different modes, referred to as *central* and *peripheral* modes (Petty and Cacioppo, 1986) or *heuristic* and *systematic* modes (Chaiken et al., 1989). On close examination, I became convinced that the difference between these “modes” lies in *contents* of the inferential rules, and that judgments reached via both modes are unexceptionally syllogistic in fact. Based on that assumption, I proceeded to examine large bodies of research interpreted as supportive of the dual mode distinction, and to offer an alternative framing of their findings cast from the lay epistemic perspective.

Persuasion Research

A pervasive finding in dual-mode inspired persuasion research has been that “peripheral” or “heuristic” cues exert judgmental impact (i.e., effect change in recipients’ attitudes or opinions) under conditions of low processing resources (e.g., where

recipients' interest in the task is low, when they are [p. 206 ↓] cognitively busy or distracted, when their need for cognition is low, etc.). By contrast, "message arguments" typically exerted their effects under high processing resources (e.g., high interest in the task, or ample cognitive capacity). Though the dual mode analyses never stated explicitly *how* message arguments or cues differentially mediate judgments, it seems plausible that in both cases the inferential mechanism is the unitary rule-based process described earlier. In case of "peripheral" or "heuristic" reasoning, the rule might be "if expert then correct," or "if supported by majority then correct," whereas in case of "central" or "systematic" reasoning the rule might be "if involving a tuition hike then unacceptable," or "if promoting individual rights then deserving of support," for example. In other words, the syllogistic rule-following process seemed the same in both presumptively distinct modes, and they only seemed to differ in contents of the rules employed.

But this fundamental commonality did little to explain the pervasively found differences in reliance on cues/heuristics or messages under different motivational or cognitive resource conditions. If the lay epistemic theory were to offer a truly comprehensive account of knowledge formation, this puzzle needed to be resolved. The way it was finally resolved rested on an empirical "discovery," involving a previously unidentified *confounding* present in many dual-mode studies. It consisted of an inadvertent linkage of low processing resources with the use of easy-to-process rules and of high processing resources with the use of relatively difficult-to-process rules (for reviews see Erb et al., 2003; Kruglanski and Gigerenzer, 2011; Kruglanski and Thompson, 1999a, 1999b; Kruglanski et al., 1999, 2007; Pierro et al., 2005).

Specifically, because in dual mode persuasion studies the message arguments were typically lengthier, more complex and placed later in the informational sequence, their processing may have imposed higher processing demands than the processing of "cues" that were invariably brief, simple, and presented upfront. When these confoundings were experimentally removed, the previously found differences between conditions under which the "cues" versus the "message arguments" (or vice versa) exerted their persuasive effects were eliminated. Once the message arguments were made shorter and simpler, they were found to mediate judgments under limited resource conditions just as did cues and heuristics in prior research. Similarly, once the cues were presented in a lengthier and more complex form (e.g., an expert's

qualifications portrayed via a lengthy resume), they mediated judgments only under ample resource conditions, as well as led to more persistent attitudes that were more predictive of behavior (Pierro et al., 2008).

Dispositional Attributions

Inattention to a potential confounding between task difficulty and informational contents afford a lay epistemic reinterpretation of findings from alternative research domains as well, unrelated to persuasion. Thus, a major attributional question concerned the process whereby an actor's behavior is causally ascribed to the *situation*, or to the actor's *disposition*. In an important paper devoted to this topic, Trope and Alfieri (1997) found that ambiguous behavior tends to be disambiguated by assimilation to the context in which it is taking place. For instance, an ambiguous facial expression is likely to be perceived as sad if the context was sad as well (e.g., a funeral), and as happy if the context was happy (e.g., a party). Once the behavior had been identified, however, and the question of its causal origin was pondered, the context plays a subtractive (rather than assimilative) role in determining the behavior's attribution. Specifically, the role of the context is subtracted to determine the role of the actor's disposition in producing the behavior. For instance, if the context was sad, an individual's sad expression would tend to not be attributed to the actor's dispositional sadness because other persons in the same situation would probably be sad as well.

[p. 207 ↓] Of present interest, Trope and Alfieri (1997) found that the assimilative process of behavior identification was independent of cognitive load, whereas the subtractive process of dispositional attribution was undermined by load. These investigators also found that invalidating the information on which the behavior identification process was based, by stating that the actor was unaware of the potential situational demands on their behavior, did not alter these identifications, whereas invalidating that same information did alter the dispositional judgments. Two alternative explanations may account for these results: (1) that the two processes are qualitatively distinct, or (2) that for some reason the behavior identification task in Trope and Alfieri's (1997) work was less demanding than the dispositional attribution task. Consistent with the latter interpretation, Trope and Gaunt (2000) discovered that when demands associated with the dispositional attribution task were lowered (e.g., by increasing

the salience of the information given), the subtraction of context from dispositional attributions was no longer affected by load. Furthermore, Chun et al. (2002) found that when the behavior identification task was made more difficult (e.g., by decreasing the salience of the information given) it was also undermined by load. Under those conditions too, invalidating the information on which the behavioral identifications were based did alter these identifications. These findings are consistent with the lay epistemic notion that, when a given judgmental task (e.g., “behavior identification” or “dispositional attribution”) is sufficiently demanding, it is exigent of cognitive resources and can be undermined by load. As Chun et al. (2002) point out, *behavior identification* and *dispositional attribution* appear to be deduced in the same (syllogistic) way from relevant evidence (Kruglanski and Gigerenzer, 2011). Furthermore, each inferential task may be more or less difficult depending on the circumstances, requiring respectively more or less resources.

Base-Rate Neglect

Inattention to the psychological difficulty inherent in using different types of evidence applies to a classic phenomenon in judgment and decision making, the neglect of base rates in favor of heuristic evidence concerning representativeness. In the original demonstrations of this effect (Kahneman and Tversky, 1973) base rates of some pertinent categories (the proportion of lawyers and engineers) in a given sample was presented to participants along with a vignette describing an individual randomly drawn from that sample. A careful analysis of typical studies on this topic carried out by Chun and Kruglanski (2006) revealed that the base rate information in many of those studies was typically presented briefly, via a single sentence, and upfront. By contrast, the individuating ('representativeness') information was presented subsequently via a relatively lengthy vignette. If one assumes that participants in such studies had sufficient motivation and cognitive capacity to process the entire informational “package” with which they were presented – they might have been challenged to fully process the later, lengthier, and hence more demanding vignette information and to have given it considerable weight in the ultimate judgment. This is analogous to the finding in persuasion studies that the lengthier, later-appearing message argument information but not the brief, upfront-appearing “cue” information

has impact under ample processing resources (e.g., of high processing motivation and cognitive capacity). If the above is true, we should be able to “move” base-rate neglect around by reversing the relative length and ordinal position in the informational sequence of the base-rate and the individuating (‘representativeness’) information. A series of studies by Chun and Kruglanski (2006) accomplished just that, and showed that when the representativeness information is rendered easier to process than the base rate information (by making the former brief and simple and the latter lengthy and complex), the representativeness information [p. 208 ↓] is relied upon more when the participants resources are depleted by cognitive load. Of even greater interest, when the base rate information is presented briefly and simply and the representativeness information is more difficult to process as in the standard paradigm, it tends to be utilized more under depleted (versus intact) resource conditions. In short, *base rates* and *representativeness* information seem to constitute different contents of evidence for likelihood estimates (e.g., that a randomly drawn individual from a population is a member of a given profession), but they afford the relevant judgments via the same inferential process. As with the persuasion and attribution tasks discussed earlier, these types of evidence may be presented in different formats varying in their difficulty of application hence requiring corresponding amounts of processing resources (Kruglanski and Gigerenzer, 2010).

Relative Relevance, Task Demands, and Processing Resources

Let us switch now to a very different confounding, namely of information type and subjective relevance of information. Often, the different types of information presented to research participants have (inadvertently) differed in their *subjective relevance* to these persons. For instance, in the domain of persuasion Pierro et al. (2004) carried out an extensive content analysis of experimental materials in persuasion studies to conclude that, typically, the “cues” presented to participants were judged as less relevant to the judgmental (attitudinal) topic than the “message arguments.” Recall that in much persuasion research the “cues” but not the “message arguments” exerted judgmental impact under low processing resources, whereas the “message arguments”

did so under high processing resources. From the present perspective, it is possible to generalize these findings in terms of the following derivations: (1) given ample processing resources, the *more relevant* information (e.g., the “message arguments” in much persuasion research) would have a greater judgmental impact than the *less relevant* information; however (2) given limited processing resources (relative to the task demands), the *easier-to-process* information (of above threshold relevance) would have a greater judgmental impact than the *difficult-to-process* information (Kruglanski and Gigerenzer, 2010).

Pierro and colleagues (2004) tested these notions in three experimental studies and found that when later, and hence the more difficult-to-process information (whether “heuristic or message argument-based”) was more subjectively relevant to the judgmental topic than early information, it exerted judgmental (persuasive impact) only under high motivation conditions but not under low motivation. By contrast, the early, less-relevant information exerted its effect only under low motivation but not under high motivation. A very different pattern obtained where the early information was more relevant than the latter information. Here, the impact of the early information invariably overrode that of the latter information: Under low processing motivation this may have been because the earlier information was easier to process than the latter information, and under high processing motivation because the early information was in fact more relevant than the latter information.

To summarize, a growing body of evidence from a variety of domains (persuasion, attribution, judgment under uncertainty, person perception) is consistent with the lay epistemic notion that the inferential mechanism whereby judgments are made is the same across different content domains of judgment. The often observed differences in the use of different kinds of information (e.g., “central” versus “peripheral” information, base rate versus heuristic information, etc) seem explicable in terms of (locally determined) difficulty or ease with which the different types of information were presented to research participants, requiring correspondingly different amounts of processing resources. Because difficulty or ease of information processing lies on a continuum [p. 209 ↓] (rather than representing a dichotomy), I concluded that the notion of two qualitatively distinct modes of making judgments is unnecessary. Instead, it seemed important to highlight the universal inferential mechanism (the syllogistic “if then” reasoning structure elaborated earlier) involved in all judgments, whose

application may be more or less difficult across different situations requiring different amounts of processing (cognitive and motivational) resources.

From Dual Modes to Dual Systems

The *dual mode* theories considered earlier (e.g., the dichotomous persuasion, or attribution models) didn't elaborate much on the process whereby each of the modes is assumed to mediate judgments, focusing instead on informational contents (e.g., of base rates versus representativeness information). In recent years, however, a different family of dualistic models was proposed in which an attempt was made to be more explicit about processes. This category of theories, generally known as *dual systems* models distinguishes between *associative* and *rule-following* processes (e.g., Kahneman, 2003; Sloman, 1996; Strack and Deutsch, 2004). Associative processes were often characterized as automatic, unconscious, and independent of resources; rule-following processes, in contrast, are typically thought of as deliberative, conscious, and resource intensive. Because of the centrality accorded in the lay epistemic theory to inferential (if-then) rules it was generally classified in the rule-following category (cf. Strack and Deutsch, 2004). If valid, the dual systems distinction would pose a challenge to the presumed universality of the knowledge formation process depicted in lay epistemic theory. Thus, the next phase of lay epistemic theory consisted in a careful examination of the dual systems claims.

Associative Processes are Rule-Based

Consider classical conditioning. Though it has been viewed as a prototypic case of an associative process, compelling evidence exists (Holyoak et al., 1989; Rescorla, 1985; Rescorla and Holland, 1982; Rescorla and Wagner, 1972) that, in fact, it is fundamentally *rule-based*. In this vein, Holyoak and colleagues (1989) explicitly stated that “representations of the environment take the form of … (if then) rules … the rat's knowledge about the relation between tones and shocks might be informally represented by a rule ‘if a tone sounds in the chamber then a shock will occur,’” (Holyoak et al., 1989: 320). In present terms, the rule represents a major

premise, that when combined with a minor premise “a tone is sounding,” warrants the deduction “a shock is coming” and elicits the warranted crouching behavior.

Whereas the work reviewed by Holyoak et al. (1989) addressed classical conditioning, a recent review of *evaluative conditioning* attests that it too is “propositional,” hence *rule-following* (Mitchell et al., *in press*). In evaluative conditioning a neutral CS (e.g., a book) is presented concomitantly with an affectively laden UCS (e.g., a smiling, or a pouting face); subsequently, the CS acquires the affective valence of the UCS. Though evaluative conditioning differs in a number of respects from classical conditioning (for discussions see Baeyens et al., 1988; Walther et al., 2005) the rule-following nature of the conditioning process appears common to both (Mitchell et al., *in press*). In short, the two paradigmatic examples of associative processes (classical and evaluative conditioning) appear to be rule-driven. This realization effectively undermines the distinction proposed by some authors between *rule-following* and *associative* processes (e.g., Gawronski and Bodenhausen, 2006; Sloman, 1996).

Automaticity of Rule-Following

Rules involved in conditioning may be applied with considerable ease and rapidity. The notion that “automatic” phenomena in the realm of skill acquisition involve a routinization of “if … then” sequences has been central to Anderson’s (1983) ACT* model that Smith (1984, 1989; Smith and Branscombe, [p. 210 ↓] 1988; Smith et al., 1988) generalized to social judgment. That research has demonstrated that social judgments represent a special case of procedural learning based on practice that strengthens the “if… then” components resulting in increased efficiency (or “automaticity”).

Awareness

Efficiency implies, in turn, a lowered need to commit attentional resources to the execution of social judgments. In William James’ (1890: 496) eloquent turn of phrase, “consciousness deserts all processes when it can no longer be of use.” According to James’ *parsimony principle* of consciousness, routinization removes the need for conscious control, rendering awareness of the process superfluous. In a related vein,

Logan (1992) suggested that automatization of certain skills effects an attentional shift to higher organizational levels.

It is in this sense, then, that some judgmental phenomena, mediated by well routinized “if...then” rules, may take place outside conscious awareness. Already Helmholtz (1910/2000) discussed the notion of unconscious *inference* in the realm of perception. More recently, social cognitive work on spontaneous trait inferences (Newman and Uleman, 1989; Uleman, 1987) suggests that lawful (i.e., rule-following) inferences presumably can occur without explicit inferential intentions, and without conscious awareness of making an inference. The spontaneous trait inference that John is “clumsy” on basis of the information that he “stepped on Stephanie's foot while dancing” (Newman and Uleman, 1989: 156), surely requires the inference rule, “if stepping on a partner's foot, then clumsy” or some variant thereof. A person who rejects that premise would be unlikely to reach that conclusion.

In summary, a variety of evidence and theoretical considerations converge on the lay epistemic notion that judgments are rule-based and in this sense, derived from “evidence.” To make a judgment is to go beyond the “information given” (Bartlett, 1932; Bruner, 1973), by using it as testimony for a conclusion in accordance with an “if... then” statement to which the individual subscribes. Such implicational structure appears fundamental to explicit human inferences (Anderson, 1983), implicit conclusion-drawing (Newman and Uleman, 1989), conditioning responses in animal learning studies (Holyoak et al., 1989; Rescorla and Wagner, 1972), and perceptual judgments of everyday objects (Gregory, 1997; Pizlo, 2001; Rock, 1983). The elementary “if... then” form appears essential to all such inferences, whether conscious or non-conscious, instantaneous or delayed, innate or learned. It is a fundamental building block from which all knowledge appears to be constructed.

Parametric Determinants of Informational Impact

Subjective relevance The lay epistemic analysis of the dual mode and dual systems theories allowed the identification of major factors involved in determining whether

the “information given,” (i.e., present in the environment or provided by a source) will influence an individual's judgments. To highlight their continuous nature, I labeled these factors judgmental *parameters*. The first parameter refers to the degree to which an individual subscribes (whether consciously or unconsciously) to a given inferential rule or major premise (“if X then Y”). The stronger the subjective implication of Y by X, the greater the evidential relevance of X to Y, and all else being equal, the stronger the tendency to reach the judgment Y upon confronting X.

It is useful to conceptually distinguish between *potential relevance* of X to Y reflecting the degree to which the “If X then Y” inferential rule has been generally learned and believed in, and *contextual or perceived relevance* reflecting the degree to which X is recognized as relevant to Y *in a given situation*. Beyond one's degree of belief in the rule, or availability of the rule in a person's memory, *perceived relevance* is affected by the rule's accessibility, difficulty of [p. 211 ↓] identifying the X, and individual's motivational and cognitive resources available for overcoming the difficulty (for discussion see Kruglanski et al., *in press*).

Difficulty of the inferential task Retrieval of the inferential rule, or the major premise, from a person's memory could be more or less difficult depending on the rule's history of activation as defined by its frequency and recency (Higgins, 1996). Similarly, situational discernment of the minor premise; that is, of information that instantiates the antecedent (X) of the rule, may be more or less difficult in different task environments. Jointly, challenges posed by activities of *retrieval* and *discernment* define the (continuous) parameter of task difficulty on which different inferential contexts may vary.⁶

Availability of cognitive resources In approaching an inferential task individuals may come to it with different reserves of cognitive resources. This may depend on their prior engagement in resource-depleting activities, or the presence of concomitant attention-demanding tasks. In general, individuals' ability to cope with difficult judgmental tasks depends on the resources at their disposal. For instance, if the discernment of relevant information is difficult to accomplish, only individuals with sufficient resources would base their judgments on that information. Similarly, if individuals' resources are depleted, they may tend to base their judgments on easy-to-process (e.g., simple, brief, and/or salient) information.

Motivation Besides depleting alternative activities, availability of attentional resources may depend on individuals' motivational states. For instance, under high (versus low) motivational involvement in a judgmental topic, individuals may mobilize greater attentional resources for an inferential task, and hence be capable of coping with relatively difficult information-processing requirements, (posed, for example, by lengthy and complex information) (for reviews see Chaiken et al., 1989; Kruglanski et al., 2007; Petty and Cacioppo, 1986).

In summary, application of the lay epistemic theory to the realm of dual mode and dual systems models resulted in a reanalysis of research findings adduced in their support, and a parametric reinterpretation focused on the interaction between the subjective relevance of evidence, the difficulty of its discernment in given circumstances, and the cognitive and motivational resources that individuals may bring to bear on the judgmental task at hand.

Epistemic Authority: Source as Evidence

According to lay epistemic theory, all evidence types function in the knowledge construction process in the same, syllogistic way. Because people's concerns typically extend beyond their domains of expertise, they often rely on other people as knowledge providers. Indeed, an essential aspect of people's social nature is their informational interdependence (Kelley and Thibaut, 1969), and the fact that they share their view of reality with significant others (Hardin and Higgins, 1996). In pondering this issue, it appeared to me that the role of *trusted sources of information* merits its own consideration in a theory of knowledge formation. Accordingly, I introduced the concept of *epistemic authority* to denote the pervasive and often determinative function of others' opinions on individuals' own judgments. Specifically, individuals may subscribe to general "if X then Y" rules in which the antecedent X denotes a given epistemic authority; for example, of an expert ("If Expert says so then it is correct"), the group ("If the Group believes so, then it is correct"), or the self ("If I believe X, then X is correct").

The concept of "epistemic authority" is akin to the notion of *source credibility* (encompassing a conjunction of perceived [p. 212 ↓] expertise and *trustworthiness*), and it addresses the extent to which an individual is inclined to treat a source's opinions

as incontrovertible evidence for her or his own judgments. The ascribed epistemic authority of various sources in the individuals' social environments may vary, and the authority of a given source may vary across domains as well as across individuals' lifespan phases.

In the social psychological literature, source characteristics (such as expertise) were often implied to offer a somewhat *inferior* counsel as to correct judgments, and were treated as suboptimal heuristics employed under limited resource conditions and when one's "sufficiency threshold" of required confidence was low (Chaiken et al., 1989). Yet, as common observations suggest, the compelling power of source authority can be considerable, (e.g., authority accorded to a religious prophet, a parent, a political leader, or the printed word) and carry substantial weight in instilling confidence in judgments. Indeed, it can often override other types of information and exert a determinative influence on individuals' opinions and corresponding behaviors.

Whereas prior treatments of source credibility affects exclusively addressed sources *external* to self (cf. Chaiken et al., 1989; Hovland et al., 1953; Kruglanski and Thompson, 1999a, 1999b; Petty and Cacioppo, 1986), the lay epistemic theory considers the *self* as a particularly important target of epistemic authority assignments. Research summarized by Kruglanski et al. (2005) has revealed: (1) developmental trends involving a decline in authority assigned to the primary caregivers, coupled with an increase in epistemic authority attributed to the *self* and involving an increase in differentiation and specificity of epistemic authorities across domains; (2) stable individual differences in epistemic authority effects; (3) a hierarchical structure and operation of epistemic authorities; and (4) the relative role of the self and external sources as perceived epistemic authorities.

Developmental Trends

Raviv et al. (1990) found that from age four to ten, (a) the perception of parents as epistemic authorities remains relatively stable, with decreases in a few knowledge areas, (b) the perception of the teacher as an epistemic authority remains stable with an increase in the area of science, and (c) the perceived epistemic authority of friends increases in the social domain. Raviv et al. (1990) also found that across age groups

the perception of teachers and friends varied more as a function of knowledge areas than the perception of parents. Specifically, parents tended to be perceived as *overall authorities*, possibly due to continued emotional and material dependence on them inducing a motivation to view them as all powerful and knowledgeable.

Individual Differences in the Distribution of Epistemic Authority Assignments across Sources

Individuals differ systematically in their distributional profiles of epistemic authority across sources. These differences, in turn, affect individuals' search for, and use of, information. Specifically, Bar (1983, 1999) found that individuals turn first to information provided by sources whom they regard as highest in epistemic authority, that they process such information more extensively, that they derive from it greater confidence, and that they tend more to act in accordance with its perceived implications.

Self-Ascribed Epistemic Authority and Learning from Experience

The concept of “experience” has long been privileged in psychological theory. The use of experiential learning in training and education has been inspired by John Dewey's (1916, 1958) instructional philosophy, Carl Rogers' (1951, 1967) person-centered approach to therapy and humanistic psychology more generally (e.g., Shafer, 1978). In social psychology, Fazio and Zanna (1981) suggested that attitudes acquired via direct experience with the attitude object are the [p. 213 ↓] strongest and most tightly related to behavior. Yet, these authors also hinted at *moderators* that qualify the power of experience in shaping attitudes. As they put it, “An attitude formed by indirect means could conceivably also be held with extreme confidence. For example, a child's attitude may be held with great confidence, even though formed indirectly because of his or her parents' extreme credibility” (Fazio and Zanna, 1981: 184).

Whether an individual would treat her or his personal experience as a reliable knowledge source may depend on this person's self-ascribed epistemic authority in a domain. In a study designed to investigate these notions, Ellis and Kruglanski (1992) assessed their participants' self-ascribed epistemic authority in mathematics via a questionnaire specifically designed for this purpose. Participants also responded to the numerical aptitude test (Cattell and Epstein, 1975) to serve as a control measure for their actual math ability, and they filled out a postexperimental questionnaire designed to assess their perceptions of their own and the instructor's epistemic authority in mathematics.

In the *experiential* condition, participants were given self-instruction booklets with exercises related to the five arithmetic rules. In the *instructional-principles* condition, the experimenter was introduced as a PhD in mathematics, and he explicitly articulated the relevant mathematical principles. In the intermediate, *instructional-examples* condition, the instructor solved the problems on the board *and* stated the arithmetic principle underlying each solution. The results indicated that method of instruction significantly interacted with participants' self-ascribed epistemic authority (SAEA). Controlling for participants' actual mathematical ability, in the *experiential* condition, participants with a high SAEA did significantly better than participants with a low SAEA. By contrast, in the *instructional principles* condition, low SAEA participants did better than their high SAEA counterparts, and in the intermediate, instructional-examples condition the high and low SAEA participants did not differ in their performance.

The foregoing findings identify a significant boundary condition on the efficacy of *experience* as a basis of *learning*. It appears that in order to be able to learn from their own experience; that is, without instruction from others, individuals need to believe in their ability to draw inferences from the experience; that is, possess high self-ascribed epistemic authority in a domain. Crucially, self-ascribed epistemic authority can be empirically distinct from actual ability in a domain. In the Ellis and Kruglanski (1992) study, the correlation between the two, though significant, was relatively low ($r = 0.36$), and the interaction between SAEA and method of instruction remained significant, even after controlling for actual mathematical ability.

Summary

Even though all evidence may function in the same (syllogistic) manner, the evidence category subsumed under the notion of epistemic authority is special. It embodies the fundamental notion that human knowledge is socially constructed and that it is heavily influenced by the opinions of significant others whose judgments one holds in high regard.

Need for Closure Research: On Epistemic Dynamics

The lay epistemic treatment of the concept of *evidence* (addressed in research programs on the *unimodel* and on *epistemic authority*) touches on the mechanism of knowledge formation, or the *how* of individuals construction of their judgments and opinions. But any knowledge is potentially insecure. New facts may become revealed, trusted epistemic sources may express views at odds with one's prior opinions, and novel hypotheses may be generated to account for familiar phenomena. Philosophers and historians of science (like Karl Popper or Thomas Kuhn), whose writings have fascinated me, highlight [p. 214 ↓] in their analyses the dynamic aspect of scientific knowledge represented by scientific revolutions (Popper, 1959), or the replacement of ruling paradigms by alternative frameworks (Kuhn, 1962).

In considering the potentially shifting aspect of human knowledge, it seemed that a major issue in this regard concerns people's open-mindedness and their inclination to continue the hypothesis generation and testing process that results in knowledge. Because the quest for new information and the generation of new hypotheses is potentially interminable, it seemed to me that the determinants of open and closed mindedness must be *subjectively driven*, and are likely to be *motivational*. In this vein, I introduced the concept of need for closure⁷ as a major psychological determinant of epistemic stability and change, inspiring the most extensive research program to date carried out within the lay epistemic framework.

I defined the need for cognitive closure as the desire for certainty on a topic. I further postulated that the magnitude of this need is determined by the perceived benefits of closure, and by the costs of lacking closure. In those terms, the need for closure was assumed to be elevated where action was required because the launching of intelligible action requires prior closure. Additionally, the need for closure was assumed to be elevated in circumstances where the possession of closure would obviate costly or laborious information processing, as may occur under time pressure, in the presence of ambient noise, or when a person is fatigued or intoxicated (see Kruglanski, 2004, for a review). In addition to the transient situational determinants of the need for closure, this motivation was also assumed to represent a dimension of individual differences, and we constructed a scale to assess it (Webster and Kruglanski, 1994). To date, this scale has been translated into numerous languages and has been shown to yield similar results with situational manipulations of the need for closure;⁸ an improved version of the scale was recently published by Roets and van Hiel (2007), and a shortened yet unpublished version of the scale has been effectively used as well (Pierro and Kruglanski, 2008).

Intrapersonal Phenomena: Seizing and Freezing

The dynamic aspect of knowledge involves what Kurt Lewin labeled as the “freezing,” “unfreezing,” and “refreezing” of individual's beliefs. The need for cognitive closure is seen to underlie these fundamental epistemic phenomena. Specifically, a heightened need for cognitive closure may induce in individuals the tendency to “seize” on early, closure-affording evidence and “freeze” upon the judgments it suggests. These tendencies were studied in reference to several classic effects in social cognition and perception, including primacy and recency effects in impression formation, the use of stereotypes, and anchoring effects among others (for reviews see Kruglanski and Webster, 1996; Kruglanski, 2004). In all these cases, a heightened need for closure induced a freezing on judgments afforded by recently or frequently activated information (e.g., early information about a social target, or a prevalent social stereotype).

In a recent demonstration of need for closure's impact on the use of recently activated information, Pierro and Kruglanski (2008) conducted a study on the *transference effect* in social judgment. The Freudian concept of transference refers to the process by which a psychotherapeutic patient superimposes onto the therapist her or his childhood fantasies about a significant childhood figure (typically a parent). Andersen and her colleagues (e.g., Andersen and Cole, 1990; Andersen et al., 1995) showed, however, that the transference effect could be part and parcel of normal social-cognitive functioning in which a significant other's schema is mistakenly applied to a new target that resembles the significant other in some respects. In a first session of Pierro and Kruglanski's (2008) experiment, participants completed the revised 14-item need for closure scale (Pierro and Kruglanski, 2005) and were asked to [p. 215 ↓] visualize and describe a significant other. In a second session, participants were presented with information about a target person with whom they expected to interact. The target person was either described in similar terms as their significant other, or was depicted as dissimilar from that person. After having studied this information, participants were presented with a recognition test of their memory for the target. Items about the target person that were not presented in the description were included in the recognition test. The degree of transference was operationally defined as the proportion of statements falsely recognized as having been included in the description of the target person that were consistent with the representation of the significant other provided in the first session. The results indicated that participants high on the need for closure exhibited a more pronounced transference effect, as indicated by higher false alarm rates, in the similar (versus dissimilar) condition than did participants low on the need for closure.

Interpersonal Phenomena

Beyond its effects on intrapersonal phenomena in the domain of social judgment, need for closure was shown to exert a variety of interpersonal phenomena in realms of linguistic expression, communication and persuasion, empathy, and negotiation behavior.

Linguistic expression Several studies looked at need for closure effects on language abstractness in interpersonal communications. Abstract language indicates a permanence of judgments across situations, and hence a greater stability of closure.

For instance, characterizing an individual's behavior in a given situation as reflecting this person's aggressiveness (an abstract depiction) implies that he or she may be expected to behave aggressively in other contexts as well. By contrast, depicting the same behavior as a "push" (i.e., a concretely situated occurrence) carries fewer trans-situational implications. Accordingly, it is possible to predict that individuals under high (versus low) need for closure would generally tend to employ abstract terms in their communications. Evidence consistent with this prediction was obtained by Boudreau et al. (1992), Rubini and Kruglanski (1997) and Webster et al. (1997).

Persuasion Research by Kruglanski et al. (1993) explored the conditions under which need for closure may increase or decrease the susceptibility to persuasion. To do this, participants were presented with information about a legal case, allowed time to process the information, and then later interacted with a partner (a fellow "juror") in order to reach a joint verdict in the case. When participants were given complete information about the case, including a (fictitious) legal analysis suggesting the appropriate verdict, individuals high (versus low) on the need for closure were less likely to be persuaded by their fellow juror (who argued for the opposite verdict). However, when high-need-for-closure individuals were given incomplete information lacking the legal analysis, they were more likely to be persuaded by their fellow juror than their low-need-for-closure counterparts. In short, individuals high (versus low) on the need for closure tend to resist persuasion attempts when they have formed a crystallized opinion about a topic, but tend to change their attitudes when presented with persuasive appeals when they lack an opinion about the topic.

Empathy Because high-need-for-closure individuals tend to "freeze" on their own perspective, they are less able to empathize with their interaction partners, especially when those are dissimilar from themselves (Webster-Nelson et al., 2003). In related work, Shytenberg et al. (2008) found that high (versus low) scorers on the need for closure scale were less sensitive to injustice done to their teammate by the experimenter (perceiving the experimenter as less unfair). Richter and Kruglanski (1999) found that individuals with high (versus low) dispositional need for closure tended less to "tune" their messages [p. 216 ↓] to their audience's perspective;⁹ as a consequence, their communications were less effectively decoded by their recipients.

Negotiation behavior Need for closure was found to exert significant effects on negotiation behavior. Thus, DeDreu et al. (1999) measured participants dispositional need for closure and then (after a 30 minute delay) had them engage in a task in which they acted as sellers and interacted with presumed buyers (actually simulated by pre-programmed responses). It was found that individuals with high (versus low) dispositional need for closure tended more to adhere to cut-off values (of minimal acceptable profits) given them by the experimenter; they also made smaller concessions to their negotiation partners and engaged in less systematic information processing. In another study on negotiation, De Dreu and Koole (1997) found that, under conditions where a majority suggested a competitive strategy, lowering participants' need for closure decreased their tendency to behave competitively and to reach an impasse.

Group Centrism

Some people are more group-oriented than others, and most people are more group-oriented in some situations than in other situations. Kruglanski et al. (2006) defined the concept of "group centrism" by the degree to which individuals strive to enhance the "groupness" of their collectivity. Groupness, in turn, has been defined by a firm, consensually supported "shared reality" (Hardin and Higgins, 1996), unperturbed by dissents and disagreements. While reality sharing has been regarded as the defining essence of groupness (e.g., Bar-Tal, 1990, 2000), its attainment may be facilitated by several aspects of group interaction, all essentially enhanced by the need for closure. At the initial phases of group formation, this can involve members' attempts to arrive at a speedy consensus, by exerting uniformity pressures on each other (DeGrada et al., 1999).

The positive relation between need for closure and autocracy (De Grada et al., 1999; Pierro et al., 2003) is consistent with Gelfand's (2008) cross-cultural research in 35 countries across the globe in which she finds a significant relationship between the country's degree of autocracy and situational constraints and the inhabitants' need for closure. Though these results may reflect the notion that high need for closure individuals tend to construct autocratic societies, they may also mean that life in tight autocratic societies tends to engender members with a high need for closure. These two

are not necessarily incompatible. Their existence and interrelation could be profitably probed in further research.

In addition to influencing group structure, intensified quest for uniformity under heightened need for closure was found to lead to an intolerance of diversity (Kruglanski et al., 2002; Shah et al., 1998). Diversity may impede the arrival at consensus, thereby reducing the group's ability to reach closure. In this vein, heightened need for closure, through the implementation of time pressure and ambient noise, has been shown to lead to a rejection of opinion deviates in a working group (Kruglanski and Webster, 1991). Elevated need for closure was also found to foster favoritism toward one's ingroup, in direct proportion to its degree of homogeneity and opinion uniformity. Finally, need for closure was found to foster outgroup derogation (Kruglanski et al., 2002; Shah et al., 1998), whose degree was *inversely* related to the outgroup's homogeneity and opinion uniformity (Kruglanski et al., 2002). These findings are consistent with the notion that high need for closure individuals are attracted to groups (whether ingroups or outgroups) that promise to afford firm shared realities to their members, providing stable cognitive closure.¹⁰

The quest for stable shared reality on part of individuals with high need for closure may express itself in conservatism and the upholding of group norms and traditions. Indeed, both political conservatism [p. 217 ↓] (Jost et al., 2003a, 2003b) and the tendency to maintain stable group norms across generational cycles (Livi, 2003) were found to be related to a heightened need for closure.

Kosic et al. (2004) found evidence that need for closure augments loyalty to one's ingroup and instills a reluctance to abandon it and 'defect' to alternative collectivities. Such loyalty persists to the extent that one's ingroup is salient in the individuals' social environment. If, however, an alternative group's views became overridingly salient, high need for closure may prompt members instead to switch groups. In this vein, Croat and Polish immigrants to Italy who were high (versus low) on need for closure tended to assimilate less to the Italian culture (i.e., they maintained loyalty to their culture of origin) if their social environment at entry consisted of their coethnics. However, if it consisted of members of the host culture (i.e., of Italians) high (versus low) need for closure immigrants tended more to "defect" and assimilate to the Italian culture.

In summary, considerable research findings attest to the considerable role that need for cognitive closure plays in intrapersonal, interpersonal, and group phenomena. At the individual level, these processes affect the formation of social judgments, attitudes and impressions. At the interpersonal level they enter into communication and persuasion, empathy, and negotiation behavior, and at the group level, into the formation of consensus and the forging of stable social realities for the members. In all these domains, and at all these levels of analysis, the need for closure has been shown to constitute a variable with implications for major classes of social psychological phenomena.

Real-World Implications

The lay epistemic theory has a variety of real world implications. The present emphasis on the concept of evidence and its subjective nature suggests that effective persuasion should be tailored to recipients' idiosyncratic inference rules, and include the provision of evidence that fits those rules. This suggests the importance of identifying the contents of those inference rules as these may differ across individuals, groups, and cultures. For instance, in recent work with detainees suspected of jihadist terrorism, an attempt has been under way to get a handle on the evidential basis of their ideology in order to counter it effectively in various national deradicalization programs (Bin Kader, 2009; Kruglanski and Gelfand, 2009). Insofar as jihadist suspects are dispersed across detention centers around the globe, and belong to diverse national and cultural groups, their beliefs and rules of inference might vary widely. Hence it is of great importance to divine what these beliefs are in order for them to be appropriately addressed. Research on these issues is currently underway in our labs.

A related topic concerns revered epistemic authorities as a basic evidential category for one's beliefs. Bar (1983, 1999) demonstrated how individuals differ in their hierarchy of epistemic authorities, and how such a hierarchy determines their readiness to accept the statements of various authorities as evidence for their own beliefs. Determining the relative epistemic authority of various sources is imperative when it comes to real-world persuasion efforts in various domains. In the deradicalization programs mentioned earlier, it would be well to ascertain the epistemic authority of different clerics for

different detainees, in order to maximize the efficacy of the religious dialogues aimed at undermining the detainees' jihadist ideology.¹¹

Similarly, determination of therapists' epistemic authority for clients may serve to improve the outcomes of clinical treatments (Kruglanski and Jaffe, 1988; Abramson et al., 1990). The same should hold for determining the epistemic authority of teachers for students in educational settings (Ellis and Kruglanski, 1992). Though some initial research exists in these domains, additional applications are well worthy of further investigation.

[p. 218 ↓] The need-for-closure concept also has numerous real-world relevancies. In clinical therapy contexts, lowering the need for closure could promote the "unfreezing" of maladaptive ('irrational') beliefs (Beck et al., 1987; Ellis and Yeager, 1989) and refreezing more adaptive ones in their stead (Kruglanski and Jaffe, 1988). The "unfreezing" and "refreezing" processes may also be applied to alternative belief systems; for example, those in the political domain. Recently, Gayer et al. (2009) carried out three studies in which they attempted to "unfreeze" Israelis' attitudes toward the Israeli-Palestinian conflict by pointing to the potential costs of failing to reach agreement on a two-state solution with the Palestinians, likely inducing a "fear of invalidity" that would lower the need for closure on this topic (Kruglanski, 2004). "Unfreezing" in this research was assessed by research participants' readiness to expose themselves to a wide array of information sources (a broad variety of press, television, and radio programs). It was found that participants' general perception of losses (in Study 1) was significantly correlated with their "unfreezing" tendency, and that priming evidence about losses (Studies 2 and 3) augmented "unfreezing."

Whereas lowering the need for closure may induce "unfreezing," heightening this need may augment "freezing." This might have several important real-world consequences as well. One of these, already commented on, involves the adoption of conservative political attitudes (Jost et al., 2003a, 2003b). The reason for this is that political conservatism includes resistance to change and traditionalism as its central features, hence "freezing" on one's prior opinions and perspectives. Though need for closure might also induce a "freezing" on prior liberal ideologies, hence inducing a "dogmatism

of the left,” the contents of liberal ideologies are less closure promoting, which makes “dogmatism of the right” more likely than “dogmatism of the left” (Jost et al., 2003a).

Because need for closure augments one's attraction to ingroups and the derogation of outgroups, it induces a particularly aggressive response in cases of intergroup conflict. In this vein, Orehk et al. (2008) recently found in a series of studies that need for closure is correlated with support for tough antiterrorism measures (including torture of suspected terrorists, and their imprisonment in foreign lands). Need for closure was also found to be positively correlated with support for a decisive and rigid presidential candidate who was likely to provide closure and to be negatively correlated with a flexible and open-minded candidate who was unlikely to do so.

In conclusion, because the processes of knowledge formation and change are related to almost all conceivable domains of human endeavors, the lay epistemic theory may enable insights to a particularly broad range of real-world problems. The full extent of this potential has yet to be explored.

Notes

1 Years later, I discovered that what Kelley meant by “lay epistemology” was a far cry from what I took it to mean. Whereas I took it to mean the *process* of knowledge construction, Kelley intended by it the *contents* of people's beliefs and their ways of representing an array of social situations and personal relationships. Kelley's theoretical work on this topic culminated in his publication with a number of colleagues of the *Atlas of Interpersonal Situations* (Kelley et al., 2003), a volume that has hardly to do with knowledge construction issues that formed the body of my own “lay epistemics.”

2 Not necessarily because of the increased *falsification opportunity* (as explained later), but because science essentially aims at reducing the observed variability in phenomena by identifying them as special instances of general principles that they share in common. In that sense, the more general the principle, the greater the theory's explanatory power, and the greater the theory's correspondence with the quintessential aim of science, which is explanation (Popper, 1959).

3 This line of reasoning also suggests that covariation shouldn't be regarded as the exclusive evidence for causality. Thus, temporal precedence of cause over effect is another implication of the cause concept and could represent an important category of evidence for causality. Causality could also be inferred [p. 219 ↓] from statements of a general epistemic authority (see Kruglanski et al., 2005) that something is the cause of something else representing another category of evidence for causality (and other things) that could lead to confident causal attributions.

4 Popper and other philosophers of science (e.g., Paul Feyerabend or Imre Lakatos) have noted that whereas knowledge formation is guided by the ideal of truth, one can never be certain that this ideal has been realized. This implies that the concept of "knowledge" is best understood in its subjective sense, as a *belief*. This hardly implies that knowledge must be solipsistic or idiosyncratic. To the contrary, knowledge typically is socially shared, and knowledge construction (whether scientific or lay) is accomplished via a communal process (Hardin and Higgins, 1996).

5 The logic of confirming a hypothesis on the basis of evidence seems immanent in the rationale of a research design in which all alternative hypotheses are controlled for that to the researcher's mind seem capable of producing a given affect, hence (subjectively) warranting the premise that (insofar as all other possible factors are controlled for) *if and only if* the focal research hypothesis is true, a given data pattern will be observed.

6 The degree to which a rule is activable in given circumstances may depend on its degree of routinization. Some rules can be routinized to the point of "automatization" which may render their use relatively independent of processing resources. In other words, the distinction between "automatic" and "deliberative" processes may represent the (quantitative) parameter of processing difficulty rather than a qualitatively different mode of judgment.

7 In early publications (e.g., Kruglanski, 1980; Kruglanski and Freund, 1983) I used the term 'need for structure' to denote the motivational stopping mechanism of knowledge formation. However, my close friend and colleague Tory Higgins convinced me that what primarily mattered to me was the closed-mindedness induced by the motivational force I was attempting to identify rather than the structured character of the resulting knowledge. Thus, the term "need for closure" was born.

8 In a recent paper, Roets et al. (2008) argued that in addition to exerting a direct motivational effect similar to that of dispositional NFCC, situational manipulations of need for closure (via time pressure or noise) exert an effect on cognitive capacity as well as manifesting itself in deteriorated task performance.

9 This may depend also on whether the object is evaluatively ambiguous or not. In this vein, Higgins and McCann (1984) found that when the object is evaluatively ambiguous or unclear there is even more tuning in to the audience, in particular if the latter has high epistemic authority.

10 This may depend on whether the individual feels that she/he could join the outgroup or not. If she/he couldn't, a highly cohesive and consensual outgroup might elicit even greater derogation.

11 In the extensive deradicalization program in Iraq headed up by US Major General Douglas Stone, it was clear that American sources would have, if anything, a negative epistemic authority for the detainees. Consequently, this successful program (on which basis no less than 18,000 suspected terrorists have been released) relied exclusively on Iraqi sources (whether religious or secular) in work with the detainees.

References

Abramson, L.Y., Alloy, L.B. and Dykman, B.M. (1990) Cognitive theory and therapy of emotional disorders: A lay epistemic reconstruction? *Psychological Inquiry* , 1, 198–219.
http://dx.doi.org/10.1207/s15327965pli0103_2

Andersen, S.M. and Cole, S.W. (1990) 'Do I know you?' The role of significant others in general social perception . *Journal of Personality and Social Psychology* , 59, 384–399.
<http://dx.doi.org/10.1037/0022-3514.59.3.384>

Andersen, S.M., Glassman, N.S., Chen, S. and Cole, S.W. (1995). Transference in social perception: The role of chronic accessibility in significant-other representations . *Journal of Personality and Social Psychology* , 69, 41–57. <http://dx.doi.org/10.1037/0022-3514.69.1.41>

Anderson, J.R. (1983) *The Structure of Cognition* . Cambridge, MA: Harvard University Press.

Baeyens, F., Crombez, G., Van den Bergh, O. and Eelen, P. (1988) Once in contact always in contact: Evaluative conditioning in resistance to extinction . *Advances in Behavior Research and Therapy* , 10, 179–199. <http://dx.doi.org/10.1016/0146-6402%2888%2990014-8>

Bar, R. (1983) Hierarchy of epistemic authority test . Unpublished MSc thesis, Tel-Aviv University.

Bar, R. (1999) The impact of epistemic needs and authorities on judgment and decision making . Unpublished PhD dissertation, Tel Aviv University.

Bar-Tal, D. (1990) *Group Beliefs: A Conception for Analyzing Group Structure, Processes, and Behavior* . New York: Springer-Verlag. <http://dx.doi.org/10.1007/978-1-4612-3298-8>

Bar-Tal, D. (2000) *Shared Beliefs in a Society: A Social Psychological Analysis* . Thousand Oaks, CA: Sage Publications. <http://dx.doi.org/10.4135/9781452204697>

Bartlett, F.C. (1932) *Remembering: A Study in Experimental and Social Psychology* . New York: Cambridge University Press.

Beck, A.T., Shaw, B.F., Rush, A.J. and Emory, G. (1987) *Cognitive Theory of Depression* . New York: The Guilford Press.

Bem, D.J. (1972) Constructing cross-situational consistencies in behavior: Some thoughts on Aker's critique of Mischel . *Journal of Personality* , 40, 17–26. <http://dx.doi.org/10.1111/j.1467-6494.1972.tb00645.x>

Bin Kader, A.H. (Ed.) (2009) *Countering Radicalism: The Next Generation and Challenges Ahead* . Singapore: Taman Bacaan Pemuda Melayu Singapura.

Boudreau, L.A., Baron, R.M. and Oliver, P.V. (1992) Effects of expected communication target expertise and timing of set on trait use in person

description . *Journal of Personality and Social Psychology* , 18, 447–451. <http://dx.doi.org/10.1177/0146167292184008>

Bruner, J.S. (1973) Beyond the Information Given: Studies in the Psychology of Knowing . Oxford: Norton.

Cattell, R.B. and Epstein, A.R. (1975) Comprehensive Ability Battery . Champaign, IL: Institute for Personality and Ability Testing.

Chaiken, S., Liberman, A. and Eagly, A.H. (1989) Heuristic and systematic information processing within and beyond the persuasion context . In J.S. Uleman, ed. and J.A. Bargh (eds), Unintended Thought . New York: Guilford Press.

Chirumbolo, A. (2002) The relationship between need for cognitive closure and political orientation: The mediating role of authoritarianism . *Personality and Individual Differences* , 32, 603–610. <http://dx.doi.org/10.1016/S0191-8869%2801%2900062-9>

Chun, W.Y. and Kruglanski, A.W. (2006) The role of task demands and processing resources in the use of base-rate and individuating information . *Journal of Personality and Social Psychology* , 91, 205–217. <http://dx.doi.org/10.1037/0022-3514.91.2.205>

Chun, W.Y., Spiegel, S. and Kruglanski, A.W. (2002) Assimilative behavior Identification can also be resource dependent: A unimodel perspective on personal-attribution phases . *Journal of Personality and Social Psychology* , 83, 542–555. <http://dx.doi.org/10.1037/0022-3514.83.3.542>

Dechesne, M., Schultz, J.M., Kruglanski, A.W., Orehek, E. and Fishman, S. (2008) A psychology of borders: need for closure and the allure of group impermeability . Unpublished manuscript, University of Maryland.

DeDreu, C.K.V. and Koole, S.L. (1997) Motivated use of heuristics in negotiation . Unpublished raw data.

DeDreu, C.K.W., Koole, S.L. and Oldersma, F.L. (1999) On the seizing and freezing of negotiator inferences: Need for cognitive closure moderates the use of heuristics

in negotiation . *Personality and Social Psychology Bulletin* , 25, 348–362. <http://dx.doi.org/10.1177/0146167299025003007>

De Grada, E., Kruglanski, A.W., Mannetti, L. and Pierro, A. (1999) Motivated cognition and group interaction: Need for closure affects the contents and processes of collective negotiations . *Journal of Experimental Social Psychology* , 35, 346–365. <http://dx.doi.org/10.1006/jesp.1999.1376>

Dewey, J. (1916) *Democracy and Education* . New York: Free Press.

Dewey, J. (1958) *Experience and Nature* . New York: Collier Macmillan.

Ditto, P.H. and Lopez, D.F. (1992) Motivated skepticism: Use of differential decision criteria for preferred and non preferred conclusions . *Journal of Personality and Social Psychology* , 63, 568–584. <http://dx.doi.org/10.1037/0022-3514.63.4.568>

Erb, H.P., Kruglanski, A.W., Chun, W.Y., Pierro, A., Mannetti, L. and Spiegel, S. (2003) Searching for commonalities in human judgment: The parametric unimodel and its dual mode alternatives . In W. Stroebe, ed. and M. Hewstone (eds), *European Review of Social Psychology* , 14, 1–47. <http://dx.doi.org/10.1080/10463280340000009>

Ellis, S. and Kruglanski, A.W. (1992) Self as epistemic authority: Effects on experiential and instructional learning . *Social Cognition* , 10, 357–375. <http://dx.doi.org/10.1521/soco.1992.10.4.357>

Ellis, A. and Yeager, R.J. (1989) *Why Some Therapies Don't Work: The Dangers of Transpersonal Psychology* . Amherst: Prometheus Press.

Fazio, R.H. and Zanna, M.P. (1981) Direct experience and attitude-behavior consistency . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 14, 161–203. San Diego, CA: Academic Press.

Festinger, L. and Carlsmith, M.E. (1959) Cognitive consequences of forced compliance . *Journal of Abnormal Psychology* , 58, 203–210. <http://dx.doi.org/10.1037/h0041593>

Gawronski, B. and Bodenhausen, G.V. (2006) Associative and propositional processes in Evaluation: An integrative review of implicit and explicit attitude change . Psychological Bulletin , 132, 692–731. <http://dx.doi.org/10.1037/0033-2909.132.5.692>

Gayer, C., Landman, S., Halperin, E. and Bar-Tal, D. (2009) Overcoming psychological barriers to peaceful conflict resolution: The role of arguments about losses . Unpublished manuscript, Tel Aviv University.

Gelfand, M.J. (2008) Situated culture: A multilevel analysis of situational constraint across 35 nations. Paper presented at Social Psychology Winter Conference , Park City, Utah.

Gregory, R. L. (1997) Eye and Brain: The Psychology of Seeing (5th edition). Princeton: Princeton University Press.

Hardin, C.D. and Higgins, E.T. (1996) Shared reality: How social verification makes the subjective objective . In R.M. Sorrentino, ed. and E.T. Higgins (eds), Handbook of Motivation and Cognition , 3, 28–84. New York: Guilford Press.

Helmholtz, H. v. (1910/2000) Helmholtz's Treatise on Physiological Optics . Bristol: Thoemmes Press

Higgins, E.T. (1996) Knowledge activation: Accessibility, applicability and salience . In E.T. Higgins, ed. and A.W. Kruglanski (eds), Social Psychology: Handbook of Basic Principles , 133–168. New York: Guilford Press.

Holyoak, K.J., Kohl, K. and Nisbett, R.E. (1989) A theory of conditional reasoning: Inductive learning within rule-based hierarchies . Psychological Review , 96, 315–340. <http://dx.doi.org/10.1037/0033-295X.96.2.315>

Hovland, C.I., Janis, I.L and Kelley, H.H. (1953) Communication and Persuasion . Newhaven, CT: Yale University Press.

James, W. (1890) The Principles of Psychology . New York: Henry Holt and Co. <http://dx.doi.org/10.1037/11059-000>

Jost, J.T., Glaser, J., Kruglanski, A.W. and Sullaway, F.J. (2003a) Exceptions that prove the rule: using a theory of motivated social cognition to account for ideological incongruities and political anomalies . *Psychological Bulletin* , 129, 383–393. <http://dx.doi.org/10.1037/0033-2909.129.3.383>

Jost, J.T., Glaser, J., Kruglanski, A.W. and Sullaway, F.J. (2003b) Political conservatism as motivated social cognition . *Psychological Bulletin* , 129, 339–375. <http://dx.doi.org/10.1037/0033-2909.129.3.339>

Kahneman, D. (2003) A perspective on judgment and choice: Mapping bounded rationality . *American Psychologist* , 58, 697–720. <http://dx.doi.org/10.1037/0003-066X.58.9.697>

Kahneman, D. and Tversky, A. (1973) On the psychology of prediction . *Psychological Review* , 80, 237–251. <http://dx.doi.org/10.1037/h0034747>

Kelley, H.H. (1967) Attribution theory in social psychology . *Nebraska Symposium on Motivation* , 15, 192–238.

Kelley, H.H. (1971) Attribution in social interaction . In E.E. Jones, ed. , D.E. Kanause, ed. , H.H. Kelley, ed. , R.E. Nisbett, ed. , S. Valins, ed. and B. Weiner (eds), *Attribution: Perceiving the Causes of Behavior* . Morristown, NJ: General Learning Press.

Kelley, H.H., Holmes, J.G., Kerr, N.L. Reis, H.T., Rusbult, C.E. and Van Lange, P.A.M. (2003) An Atlas of Interpersonal Situations . Cambridge: Cambridge University Press.

Kelley, H.H. and Thibaut, J.W. (1969) Group problem solving . In G. Lindzey, ed. and E. Aronson (eds), *The Handbook of Social Psychology* . Reading, MA: Addison-Wesley.

Kosic, A., Kruglanski, A.W., Pierro, A. and Mannetti, L. (2004) Social cognition of immigrants acculturation: effects of the need for closure and the reference group at entry . *Journal of Personality and Social Psychology* , 86, 796–813. <http://dx.doi.org/10.1037/0022-3514.86.6.796>

Kruglanski, A.W. (1975) The endogenous-exogenous partition in attribution theory . *Psychological Review* , 82, 87–406. <http://dx.doi.org/10.1037/0033-295X.82.6.387>

Kruglanski, A.W. (1980) Lay epistemologic process and contents: Another look at attribution theory . *Psychological Review* , 87, 70–87. <http://dx.doi.org/10.1037/0033-295X.87.1.70>

Kruglanski, A.W. (2004) *The Psychology of Closed Mindedness* . New York: Psychology Press.

Kruglanski, A.W., Alon, S. and Lewis, T. (1972) Retrospective misattribution and task enjoyment . *Journal of Experimental Social Psychology* , 8, 493–501. <http://dx.doi.org/10.1016/0022-1031%2872%2990075-3>

Kruglanski, A.W. and Gigerenzer, G. (2011) Intuitive and deliberative judgments are based on the same principles . *Psychological Review* .

Kruglanski, A.W., Dechesne, Orehek, E. and Pierro, A. (in press) Three decades of lay epistemics: The why, how and who of knowledge formation . *European Review of Social Psychology* .

Kruglanski, A. and Freund, T. (1983) The freezing and un-freezing of lay-inferences: Effects on impressional primacy, ethnic stereotyping and numerical anchoring . *Journal of Experimental Social Psychology* , 19, 448–468. <http://dx.doi.org/10.1016/0022-1031%2883%2990022-7>

Kruglanski, A.W. and Gelfand, M.J. (2009) Deradicalization . Paper presented at?

Kruglanski, A.W. and Jaffe, Y. (1988) Curing by knowing: The epistemic approach to cognitive therapy . In L. Abramson (ed.) *Social Cognition and Clinical Psychology* . New York: The Guilford Press.

Kruglanski, A.W., Pierro, A., Mannetti, L. and DeGradaE. (2006) Groups as epistemic providers: Need for closure and the unfolding of group-centrism . *Psychological Review* , 113, 84–100. <http://dx.doi.org/10.1037/0033-295X.113.1.84>

Kruglanski, A.W., Pierro, A., Mannetti, L., Erb, H-P. and Chun, W.Y. (2007) On the parameters of social judgment . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 39, 345–392. <http://dx.doi.org/10.1016/S0065-2601%2806%2939005-3>

Kruglanski, A.W., Riter, A., Amitai, A., Margolin, R., Shabtai, L. and Zaksh, D. (1975a) Can money enhance intrinsic motivation: A test of the content-consequence hypothesis . *Journal of Personality and Social Psychology* , 31, 744–750. <http://dx.doi.org/10.1037/0022-3514.31.4.744>

Kruglanski, A.W., Riter, A., Arazi, D. Agassi, R., Montejo, J., Peri, I. and Peretz, M. (1975b) The effects of task intrinsic rewards on extrinsic and intrinsic motivation . *Journal of Personality and Social Psychology* , 21, 699–706. <http://dx.doi.org/10.1037/0022-3514.31.4.699>

Kruglanski, A.W., Raviv, A., Bar-Tel, D., Raviv, A., Sharvit, K., Ellis, S. Bar, R., Pierro, A. and Mannetti, L. (2005) Says who? Epistemic authority effects in social judgment . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 37, 345–392. <http://dx.doi.org/10.1016/S0065-2601%2805%2937006-7>

Kruglanski, A.W., Shah, J.Y., Pierro, A. and Mannetti, L. (2002) When similarity breeds content: Need for closure and the allure of homogeneous and self-resembling groups . *Journal of Personality and Social Psychology* , 83, 648–662. <http://dx.doi.org/10.1037/0022-3514.83.3.648>

Kruglanski, A.W. and Thomson, E.P. (1999a) Persuasion by a single route: A view from the unimodel . *Psychological Inquiry* , 10, 83–109. <http://dx.doi.org/10.1207/S15327965PL100201>

Kruglanski, A.W. and Thomson, E.P. (1999b) The illusory second mode or, the cue is in the message . *Psychological Inquiry* , 10, 182–193 <http://dx.doi.org/10.1207/S15327965PL100217>

Kruglanski, A.W., Thomson, E.P. and Spiegel, S. (1999) Separate or equal? Bimodal notions of persuasion and a single process ‘unimodel’ . In S. Chiken, ed. and Y. Trope (eds.), *Dual-Process Theories in Social Psychology* , pp. 293–313. New York: Guilford Press.

Kruglanski, A.W. and Webster, D.M. (1991) Group members' reactions to opinion deviates and conformists at varying degrees of proximity to decision deadline and of

environmental noise . *Journal of Personality and Social Psychology* , 61, 212–225.
<http://dx.doi.org/10.1037/0022-3514.61.2.212>

Kruglanski, A.W., Webster, D.M. and Klem, A. (1993) Motivated resistance and openness to persuasion in the presence or absence of prior information . *Journal of Personality and Social Psychology* , 65, 861–876. <http://dx.doi.org/10.1037/0022-3514.65.5.861>

Kuhn, T.S. (1962) *The Structure of Scientific Revolutions* . Chicago: University of Chicago Press.

Livi, S. (2003) Il bisogno di chiusura cognitiva e la trasmissione delle norme nei piccoli gruppi [The need for cognitive closure and norm transmission in small groups]. Unpublished PhD dissertation, University of Rome ‘La Sapienza’.

Logan, G.D. (1992) Attention and preattention in theories of automaticity . *American Journal of Psychology* , 105, 317–339. <http://dx.doi.org/10.2307/1423031>

Mitchell, C.J., De Houwer, J. and Lovibond, P.F. (in press) The propositional nature of human associative learning . *Behavioral and Brain Sciences* .

Newman, L.S. and Uleman, J.S. (1989) Spontaneous trait inference . In J.S. Uleman, ed. and J.A. Bargh (eds), *Unintended Thought* . New York: Guilford Press.

Orehek, E., Fishman, S., Dechesne, M., Doosje, B., Kruglanski, A.W., Cole, A.P., et al. (2008) Certainty quest and the social response to terrorism . Unpublished manuscript, University of Maryland.

Petty, R.E. and Cacioppo, J.T. (1986) The elaboration likelihood model of persuasion . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 19, 123–205. <http://dx.doi.org/10.1016/S0065-2601%2808%2960214-2>

Pierro, A. and Kruglanski. A.W. (2005) Revised need for closure scale . Unpublished data, University of Rome ‘La Sapienza’.

- Pierro, A. and Kruglanski, A.W. (2008) 'Seizing and freezing' on a significant-person schema: Need for closure and the transference effect in social judgment . *Personality and Social Psychology Bulletin* , 34, 1492–1503. <http://dx.doi.org/10.1177/0146167208322865>
- Pierro, A. and Mannetti, L. (2004) Motivated consumer search behavior: The effects of epistemic authority . Unpublished manuscript, University of Rome 'La Sapienza'.
- Pierro, A., Mannetti, L., DeGrada, E., Livi, S. and KruglanskiA.W. (2003) Autocracy bias in groups under need for closure . *Personality and Social Psychology Bulletin* , 29, 405–417. <http://dx.doi.org/10.1177/0146167203251191>
- Pierro, A., Mannetti, L., Erb, H-P., Spiegel, S. and Kruglanski, A.W. (2005) Informational length and order of presentation as determinants of persuasion . *Journal of Experimental Social Psychology* , 41, 458–469. <http://dx.doi.org/10.1016/j.jesp.2004.09.003>
- Pierro, A., Mannetti, L., Kruglanski, A.W. and Sleeth-Keppler, D. (2004) Relevance override: On the reduced impact of 'cues' under high motivation conditions or persuasion studies . *Journal of Personality and Social Psychology* , 86, 251–264. <http://dx.doi.org/10.1037/0022-3514.86.2.251>
- Pierro, A., Mannetti, L., Orehek, E. and Kruglanski, A.W. (2008) Persistence of attitude change and attitude-behavior correspondence based on extensive processing of source information . Unpublished manuscript, University of Rome 'La Sapienza'.
- Pizlo, Z. (2001) Perception viewed as an inverse problem . *Vision Research* , 41, 3145–3161. <http://dx.doi.org/10.1016/S0042-6989%2801%2900173-0>
- Popper, K.R. (1959) *The Logic of Scientific Discovery* . New York: Basic Books.
- Raviv, A., Bar-Tal, D., Raviv, A. and Houminer, D (1990) Development in children's perception of epistemic authorities . *British Journal of Developmental Psychology* , 8, 157–169. <http://dx.doi.org/10.1111/j.2044-835X.1990.tb00830.x>

Rescorla, R.A. (1985) Conditioned inhibition and facilitation . In R.R. Miller, ed. and N.E. Spear (eds), *Information Processing as Animals: Conditioned Inhibition* , pp. 299–326. Hillsdale, NJ: Erlbaum.

Rescorla, R.A. and Holland, P.C. (1982) Behavioral studies of social learning in animals . *Annual Review of Psychology* , 33, 265–308. <http://dx.doi.org/10.1146/annurev.ps.33.020182.001405>

Rescorla, R.A. and Wagner, A.R. (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement . In A.H. Black, ed. and W.F. Prokasy (eds), *Classical Conditioning II: Current Research and Theory* , pp. 64–99. New York: Appleton-Century-Crofts.

Richter, L. and Kruglanski, A.W. (1999) Motivated search for common ground: Need for closure effects on audience design in interpersonal communication . *Personality and Social Psychology Bulletin* , 25, 1101–1114. <http://dx.doi.org/10.1177/01461672992512004>

Rock, I. (1983) *The Logic of Perception* . Cambridge, MA: MIT Press.

Roets, A. and van Hiel, A. (2007) Separating ability from need: Clarifying the dimensional structure of the Need for Closure Scale . *Personality and Social Psychology Bulletin* , 33, 266–280. <http://dx.doi.org/10.1177/0146167206294744>

Roets, A., van Hiel, A., Cornelis, I. and Soetens, B., (2008) Determinants of task performance and invested effort: A need for closure by relative cognitive capacity interaction analysis . *Personality and Social Bulletin* , 34, 779–792. <http://dx.doi.org/10.1177/0146167208315554>

Rogers, C.R. (1951) *Client-Centered Therapy: It's Current Practice, Implications, and Theory* . Oxford: Houghton Mifflin.

Rogers, C.R. (1967) *The Therapeutic Relationship and Its Impact: A Study of Psychotherapy with Schizophrenics* . Oxford: Wisconsin Press.

Rubini, M. and Kruglanski, A.W. (1997) Brief encounters ending in estrangement: Motivated language-use and interpersonal rapport . *Journal of Personality and Social Psychology* , 12, 1047–1060. <http://dx.doi.org/10.1037/0022-3514.72.5.1047>

Shytenberg, G., Gelfand, M.J., Imai, L., Bell, C.B. and Mayer, D. (2008) When others' injustices matters: The role of collectivism and epistemic needs . *Journal of Personality and Social Psychology* .

Shafer, J.B.P. (1978) *Humanistic Psychology* . Englewood Cliffs, NJ: Prentice Hall.

Shah, J.Y., Kruglanski, A.W. and Thompson, E.P. (1998) Membership has its (epistemic) rewards: Need for closure effects on ingroup bias . *Journal of Personality and Social Psychology* , 75, 383–393. <http://dx.doi.org/10.1037/0022-3514.75.2.383>

Sloman, S.A. (1996) The empirical case for two systems of reasoning . *Psychological Bulletin* , 119, 3–22. <http://dx.doi.org/10.1037/0033-2909.119.1.3>

Smith, E.R. (1984) Model of social inference processes . *Psychological Review* , 91, 392–413. <http://dx.doi.org/10.1037/0033-295X.91.3.392>

Smith, E.R. (1989) Procedural efficiency: general and specific components and effects on social judgment . *Journal of Experimental Social Psychology* , 25, 500–523. <http://dx.doi.org/10.1016/0022-1031%2889%2990003-6>

Smith, E.R. and Branscombe, N.R. (1988) Category accessibility as implicit memory . *Journal of Experimental Social Psychology* , 24, 490–504. <http://dx.doi.org/10.1016/0022-1031%2888%2990048-0>

Smith, E.R., Branscombe, N.R. and Bormann, C. (1988) Generality of the effects of practice on social judgment tasks . *Journal of Personality and Social Psychology* , 54, 385–395. <http://dx.doi.org/10.1037/0022-3514.54.3.385>

Strack, F. and Deutsch, R. (2004) Reflective and impulsive determinants of social behavior . *Personality and Social Psychology Review* , 8, 220–247. http://dx.doi.org/10.1207/s15327957pspr0803_1

TropeY. and Alfieri, T. (1997) Effortfulness and flexibility of dispositional judgment processes . *Journal of Personality and Social Psychology* , 73, 662–674. <http://dx.doi.org/10.1037/0022-3514.73.4.662>

Uleman, J.S. (1987) Consciousness and control: The case of spontaneous trait inferences . *Personality and Social Psychology Bulletin* , 13, 337–354. <http://dx.doi.org/10.1177/0146167287133004>

Walther, E., Nagengast, B. and Trasselli, C. (2005). Evaluative conditioning in social psychology: Facts and speculations . *Cognition and Emotion* , 19, 175–196. <http://dx.doi.org/10.1080/02699930441000274>

Webster, D.M. and Kruglanski, A.W. (1994) Individual differences in need for cognitive closure . *Journal of Personality and Social Psychology* , 67, 1049–1062. <http://dx.doi.org/10.1037/0022-3514.67.6.1049>

Webster, D.M., Kruglanski, A.W. and Pattison, D.A. (1997) Motivated language use in intergroup contexts: Need for closure effects on the linguistic intergroup bias . *Journal of Personality and Social Psychology* , 72, 1122–1131. <http://dx.doi.org/10.1037/0022-3514.72.5.1122>

Webster-Nelson, D., Klein, C.T. and Irvin, J. E. (2003) Motivational antecedents of empathy: Inhibiting effects of fatigue . *Basic and Applied Social Psychology* , 25, 37–50. http://dx.doi.org/10.1207/S15324834BASP2501_3

Weiner, B. (1979) A theory of motivation for some classroom experiences . *Journal of Educational Psychology* , 71-3–25. <http://dx.doi.org/10.1037/0022-0663.71.1.3>

Weiner, B. (1985) An attributional theory of achievement motivation and emotion . *Psychological Review* , 92, 548–573. <http://dx.doi.org/10.1037/0033-295X.92.4.548>

<http://dx.doi.org/10.4135/9781446249215.n11>

Handbook of Theories of Social Psychology: Volume 1

The Elaboration Likelihood Model

Contributors: Richard E. Petty & Pablo Briñol

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "The Elaboration Likelihood Model"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n12>

Print pages: 224-246

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n12>

[p. 224 ↓]

Chapter 11: The Elaboration Likelihood Model

Richard E. Petty, ed. and Pablo Briñol, ed.

Abstract

This chapter traces the development of the elaboration likelihood model (ELM) across three decades of research. The ELM began as a theory about the processes responsible for attitude change and the strength of the attitudes that result from those processes. It has now been applied to a wide variety of judgmental change phenomena. By focusing on the core mechanisms of change, the ELM has served to organize the many different theories, outcomes, and variables relevant to persuasion and influence. This review describes four fundamental ideas from the ELM and six phases of ELM research. A key strength of the ELM is that it provides a useful framework from which to understand the moderation and mediation of changes in attitudes as well as other judgmental outcomes from reducing prejudice to the impact of classic heuristics that influence choice and decision making.

Introduction

The study of attitudes and persuasion has one of the longest histories in social psychology (see Briñol & Petty, *in press*). At one point, the study of attitudes was considered to be the single most indispensable topic in the discipline (Allport, 1935). Empirical studies on persuasion were among the first in the field and Carl Hovland's massive program of research on attitude change during and after World War II set the

core topics and provided the research agenda for decades afterwards (see McGuire, 1968). The study of attitude change became so popular that by the 1970s, there were hundreds of studies and many conceptual analyses. Indeed, so much research and so many specific theories had accumulated that this area of inquiry was in danger of collapse from the weight of competing theories and conflicting findings.

One problem was that seemingly simple variables such as the credibility of the message source that were expected to have a relatively straightforward impact on persuasion according to the persuasion theories of the time, instead produced a mystifying diversity of findings. The accumulated research results just did not support the widespread simple main effect assumptions that accepted theories had for the persuasion [p. 225 ↓] outcomes of many variables (Petty, 1997). For instance, associating a message with an expert source, though usually good for persuasion, sometimes led to reduced influence. Another critical problem was that the core concept of attitudes was under attack largely because in some studies attitudes appeared to be consequential (e.g., guiding behavior) but more often, it seemed, they were not (Wicker, 1969). The surprising complexity of research findings caused most reviewers of the attitudes literature in the 1970s to be quite pessimistic about the usefulness of additional research (e.g., see Fishbein and Ajzen, 1972; Kiesler and Munson, 1975).

It was against this backdrop that the elaboration likelihood model (ELM) was born as a collaborative effort between Richard Petty and John Cacioppo while they were graduate students at Ohio State University in the mid-1970s. For his dissertation, Petty decided to tackle the problem of why some attitude changes persisted over time whereas others were very ephemeral. Drawing on the available literature and personal experience, Petty speculated that when attitude change was produced thoughtfully (such as after listening to strong arguments presented by John Cacioppo), the new judgments were relatively persistent whereas when attitude change occurred with relatively little thinking (such as when deciding you liked someone based on a first impression), the resulting judgment was more transitory. When Tim Brock, advisor for the dissertation, first learned about the planned studies on the persistence of persuasion, he challenged Petty to be more grandiose and propose a more general theory of attitude change rather than focusing on the more narrow attitude persistence hypothesis alone.

Intrigued by the challenge, Petty drew his friend and roommate, Cacioppo, into a long series of late-night (sometimes heated) conversations about the formation and change of attitudes that served as the foundation for the theory that was to come. The core two routes to persuasion idea (i.e., relatively thoughtful or not) was first presented in the final chapter of Petty's dissertation following empirical studies focusing on memory for one's own issue-relevant thoughts as a determinant of the persistence of attitude change (see Petty, 1977). The dissertation also benefited greatly by the presence of Tony Greenwald on the Ph.D. committee. Greenwald (1968) had earlier proposed a "cognitive response" approach to attitude change which focused on a high elaboration-mechanism by which persuasion occurred or was resisted (i.e., actively generating favorable or unfavorable thoughts to the message arguments). The addition of a low thinking route to persuasion built on Greenwald's earlier approach.

The two routes to persuasion theory did not receive the *elaboration likelihood model* (ELM) name until it was first used in a textbook on persuasion that Petty and Cacioppo (1981) wrote in their first few years out of graduate school. The name was developed after John Harvey, editor of the series in which the book was to appear, advised that a formal name was *essential* if the idea was to stick. In hindsight, it was clear that he was right! The title of the theory was selected to convey the core idea that the high versus low thought processes of persuasion formed a continuum rather than a discrete pair.

Interestingly, at about the same time, Shelly Chaiken, a graduate student at the University of Massachusetts working on her dissertation with Alice Eagly, was also developing the idea that persuasion was sometimes the result of effortful thinking but was sometimes the result of a lower effort reliance on simple heuristics such as "experts are correct" (see Chaiken, 1978). Without awareness of each other's dissertation work, both Petty and Chaiken entered the job market in the same year and even competed for the same positions at several universities, probably (to the bewilderment of the audience) giving similar job talks. Over time, they became good-natured rivals and friends. Chaiken's theory was first called the heuristic model (Chaiken, 1987) to emphasize this unique low effort mechanism [p. 226 ↓] of persuasion and eventually the heuristic-systematic model (HSM) in order to highlight the low versus high effort processes involved (see Chaiken et al., 1989). Although the ELM and HSM stem from different conceptual traditions (i.e., cognitive response theory versus message learning theory) and use different language and terminology, the theories have far more in

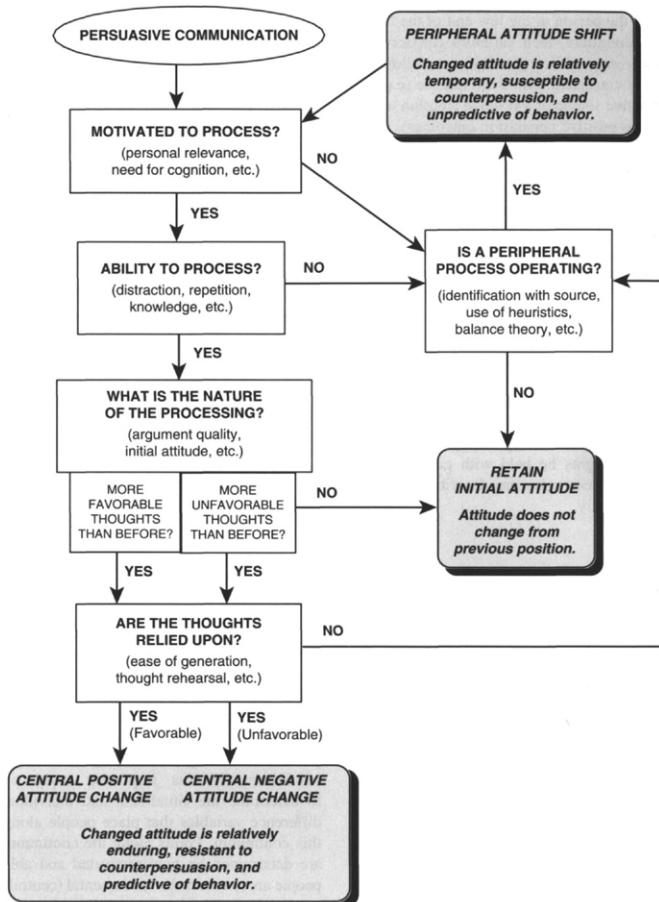
common than they have points of divergence (see Petty and Wegener, 1998, 1999). Most importantly, the joint appearance of these theories and the research inspired by them did much to foster a more general interest in what became an explosion of dual process (see Chaiken and Trope, 1999) and dual system (e.g., Deutsch and Strack, 2006) approaches to judgment.

In any case, by the mid-1980s a good number of studies had emerged testing various ELM ideas and Petty and Cacioppo (1986a) summarized the accumulated research in a monograph in which the ELM was first presented as a series of seven formal postulates (see also Petty and Cacioppo, 1986b). In the years since then, as more work on the ELM was published, various new summaries of research guided by the ELM have appeared (e.g., Petty and Wegener, 1999) of which this chapter is the most current. From its inception, the ELM was developed to account for the complicated, contradictory, and even perplexing results obtained in the accumulated persuasion literature. It also aimed to provide an integrative framework from which past research findings could be understood as well as new predictions generated in the attitudes domain and beyond. In describing the development of the ELM over time, we will also highlight some of the key people who played important roles.

Four Core Elm Ideas

The ELM has been presented both schematically (e.g., Petty, 1977; Petty and Cacioppo, 1981, 1986a, 1986b; see [Figure 11.1](#)) and as a series of formal propositions (Petty and Cacioppo, 1986a, 1986b; Petty and Wegener, 1999). Stripped to its bare bones, however, the ELM does four essential things. First, it highlights the fact that modifying people's attitudes or other judgments can be done with a high degree of thought or a relatively low degree of thought. That is, the "elaboration continuum" ranges from low to high.

Figure 11.1 Schematic depiction of the Elaboration Likelihood Model



Second, the ELM holds that there are numerous specific processes of change that operate along this continuum (e.g., classical conditioning and mere exposure require relatively little thought and operate at the low end of the continuum, but expectancy-value and cognitive response models require high degrees of thought and operate along the upper end of the continuum). When the operation of processes at the low end of the continuum determines attitudes, persuasion is said to follow the *peripheral route* whereas when the operation of processes at the high end of the continuum determines attitudes, persuasion is said to follow the *central route*. Of course, much of the time, persuasion is determined by a mixture of these processes.

The third thing the ELM does is to postulate that it matters whether persuasion occurs as the result of relatively high or low amounts of thought. This is because the degree of thought behind a judgment determines how consequential that judgment is. Specifically, the more a judgment is based on thinking about the merits of an issue, the more it tends to persist over time, resist attempts at change, and has consequences for other judgments and behavior (Petty et al., 1995).

The fourth and arguably most useful thing that the ELM does is to organize the many specific processes by which variables can affect attitudes into a finite set that operate at different points along the elaboration continuum. For example, the ELM postulates that one of the things that variables such as the attractiveness of the source of a message or the incidental emotion a person is experiencing can do is to affect how much thinking a person is doing – placing them somewhere [p. 227 ↓] [p. 228 ↓] along the elaboration continuum. However, if circumstances have already conspired to place the person at the low end of the thinking continuum, then variables can serve as simple cues, affecting attitudes in a direction that is consistent with their valence (e.g., an attractive source or a positive emotion would lead to positive persuasion outcomes). If the person is at the high end of the elaboration continuum, then there are three other ways in which the variable can affect judgments. Specifically, the variable (1) can be examined as an argument (does the fact that the source is attractive or that the person feels good provide some relevant evidence as to the true merit of what is being advocated?), (2) can affect the valence of the thoughts that come to mind (e.g., exposure to an attractive source or being in a good mood can make positive thoughts more likely to come to mind), and/or (3) can affect a structural feature of the thoughts generated (e.g., an attractive source or feeling happy could make one's thoughts be held with greater confidence). These roles are described in more detail shortly.

Six Phases of ELM Research

Given a theory with the ambitions and complexity of the ELM, it could not, of course, be tested in a single study, or two or ten. Instead, research on the theory proceeded in a series of stages, and our review will follow these phases in a roughly chronological order. The first stages of work, conducted mostly by Petty and Cacioppo and their various peer and student collaborators, focused on the four core ideas just outlined.

Thus, the first phase focused on simply establishing that there was a thinking continuum and that this continuum was consequential for persuasion. The second stage focused on providing evidence for the idea that the mechanism of persuasion could be different under high and low thinking conditions. A third phase examined the consequences of attitudes changed by high versus low thinking conditions. A fourth phase provided evidence for the so-called “multiple roles” postulate – the idea that any one persuasion variable could affect attitudes in different ways depending on the likelihood of thinking.

Once the four core ELM ideas were supported in the first phases of the research program, a fifth phase of research focused on extending the ELM principles to other judgmental areas beyond the persuasion domain. Although work on each of these phases continues, the most active current phase of research focuses on exploration of a particular role that variables can assume in modifying attitudes or other judgments. Whereas prior research focused on primary cognition – the original association of an attitude object with some attribute – current work is examining the role of secondary cognition (i.e., metacognition). In particular, this work focuses on how and when people assess the validity of their thoughts and what the consequences of this are. We next review the six phases of research on the ELM and present a study that illustrates each.

Phase 1: Exploring the Elaboration Continuum

In contrast to the earliest attitude change theories that focused on just one process of change (e.g., classical conditioning; Staats and Staats, 1958), the ELM allows for multiple processes that can involve different degrees of thinking. Because different processes of change occur along the thinking continuum, it was important early on to determine the situational and individual difference variables that place people along this continuum. Points along the continuum are determined by how motivated and able people are to assess the fundamental (central) merits of a person, issue, or position (i.e., the attitude object). The ELM assumes that when making an evaluative judgment, the default goal is to determine how good or bad the object truly is. That is, people want to have [p. 229 ↓] attitudes that are subjectively correct. However, people neither have the desire nor the ability to attain equal confidence in every attitude. Thus, motivational

and ability factors will determine how much thinking they do in any given situation. For example, it is not worthwhile to exert considerable mental effort to achieve correctness in all situations and people do not always have the requisite knowledge, time, or opportunity to thoughtfully assess the merits of a proposal.

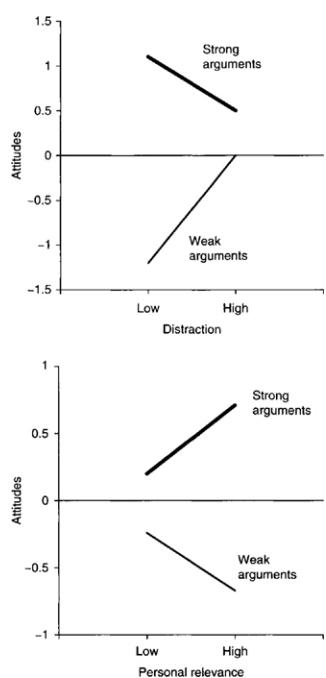
Amount of Thinking

In early research relevant to the ELM, it was useful to show that differences in the underlying extent of thinking (elaboration) could provide an explanation for the persuasive effects of variables that had been accounted for in different ways by prior theories.¹ The idea that variables could affect the extent of thinking was also important in explaining how any one variable could both increase and decrease persuasion. As an example, consider a variable like external distraction. Prior research guided by a message-learning approach (e.g., Hovland et al., 1953) suggested that distraction should be bad for persuasion because it would disrupt learning of the message arguments. Prior research guided by dissonance theory, however, suggested that distraction could be good for persuasion because people would have to justify the extra effort they put into processing the message (Baron et al., 1973). Another possibility, suggested by the elaboration continuum idea, was that distraction would affect how much thinking people did about a message.

Imagine a person who is exposed to a message containing eight cogent arguments. The high elaboration processor might think of two or three favorable implications of each of the arguments, whereas the low elaboration processor might think of only one favorable implication (because he or she is not thinking as much). The effect of this is that the high elaboration processor will likely have more favorable attitudes toward the issue than the low elaboration processor because he or she will have generated more favorable implications of the strong arguments presented and use these thoughts as a guide to the correct attitude. Thus, if distraction reduces processing, when the message contains strong arguments, distraction will be associated with reduced persuasion because fewer favorable thoughts will be generated. This result would be consistent with both learning theory and the ELM. However, what if the message contains weak rather than strong arguments? In this case the high elaboration processor might think

of many unfavorable implications of the arguments (i.e., counterarguments), whereas the low elaboration processor might think of only a few. This time the effect is that the high elaboration processor will have less favorable attitudes toward the issue than the low elaboration processor because he or she will have generated more unfavorable implications of the specious arguments presented. When this hypothesis was tested in an empirical study in which the extent of distraction and argument quality were jointly manipulated, the interaction pattern on the attitude data that was expected by the elaboration hypothesis was obtained (Petty et al., 1976, see [Figure 11.2](#), top panel). That is, distraction reduced persuasion when the arguments were strong but increased persuasion when the arguments were weak.²

Figure 11.2 Interactions of variables with argument quality. Top panel depicts distraction (data from Petty et al., 1976). Bottom panel depicts personal relevance (data from Petty and Cacioppo, 1979b)



Following this study on distraction – the first to use a manipulation of argument quality to examine how a variable affects thinking – many other investigations have also used

this paradigm. Today, there is a long list of variables that have been shown to affect the extent of thinking and thereby influence attitudes. These variables include message repetition, accountability, and emotion, to name just a few (see Petty and Wegener, 1998, for a review). Perhaps the most studied variable in this regard is the personal relevance of the communication. Linking the message to virtually any aspect of the self appears to increase motivation to think about it (Petty and Cacioppo, 1990). For example, in one study (Petty and [p. 230 ↓] Cacioppo, 1979b), undergraduate students were told that a new comprehensive exam policy was going into effect next year or not until ten years in the future. The proposal for requiring students to pass an exam in their major area of concentration as a prerequisite to graduation was supported with either strong or weak arguments. When the policy was said to affect the students personally, argument quality effects were larger than when the policy had no personal relevance (see bottom panel of [Figure 11.2](#)). Or viewed differently, increasing personal relevance tended to increase persuasion when the arguments were strong but to decrease persuasion when the arguments were weak.

In the previous examples, people were more motivated to be thoughtful if the message was linked to the self and they were more able to attain a thoughtful assessment of the arguments if the persuasion context was free of distraction. Although a motive to [p. 231 ↓] be accurate is assumed to be the default goal and underlies the effect on elaboration of variables such as personal relevance, it is not the only motive that affects the extent of information processing. For example, putting people in a positive mood gets them to think more about pleasant messages, not because positive moods or pleasant messages increase the desire to be correct, but because thinking about a pleasant message is hedonically rewarding, and people in positive moods are especially attentive to the hedonic consequences of their actions (Wegener and Petty, 1994; Wegener et al., 1995). In a similar vein, some individuals generally take greater pleasure in thinking than others and thus these individuals (i.e., those high in *need for cognition*; Cacioppo and Petty, 1982) tend to engage in effortful thought because of its intrinsic enjoyment without respect to the importance of the issue or the need to be correct (see Cacioppo et al., 1996; Petty et al., 2009b for reviews). People also generally think more when something makes them feel doubt rather than certainty in their attitudes because doubt is generally less comfortable than certainty and people can try to reduce that discomfort by reassessing their attitudes (see Briñol et al., 2006).³

Biases in Thinking

It is important to note that just because a person is thinking intently about a message, the ELM does not assume that the thinking will be totally objective. Rather, the ELM holds that variables can affect not only how much thinking a person is doing, but also whether the thoughts are relatively objective or biased. Consider the personal relevance of the message. We have seen that the more the message connects to the self, the more thinking the message elicits. But is that thinking relatively objective or biased? As outlined by Petty and Cacioppo (1990), this depends on a number of factors. For example, does the message threaten or support one's current views? If a message is relevant (versus irrelevant) to one's outcomes, values, identities, possessions, and so forth, it will engage more processing. If the message takes a position that is consistent and supportive of one's outcomes, values, and so forth, it will lead to positively biased processing. However, if it takes a position that is counter to or threatening to one's outcomes, values, and so forth, it will lead to negatively biased processing (see Petty et al., 1992).

A number of motivational and ability variables have been shown to bias processing (i.e., affect the valence of the thoughts that come to mind). For example, if a message induces psychological reactance (see Brehm, 1966) by placing undue pressure on an individual to change his or her mind, the person will be motivated to resist and therefore counterargue the message (see Petty and Cacioppo, 1979a). If balance motives (Heider, 1958) are operating, people would prefer to adopt the position of a liked source but distance themselves from a disliked source. If impression management motives (Tedeschi et al., 1971) are in ascendance, people would prefer to hold whatever position they think would be ingratiating and avoid those that would make them look bad. If self-affirmation motives (Steele, 1988) are high, people prefer the position that would make them feel best about themselves, and so forth. Clearly, there are a host of motives that can produce biases in information processing (see Briñol and Petty, 2005). Or, in the absence of any motivational forces, certain factors can uniquely enable positive or negative thoughts (e.g., positive emotions can make positive thoughts more accessible; Petty et al., 1993). In sum, the ELM holds that two of the ways in which a variable can affect attitudes are to (1) affect how much thinking takes place (amount of

thinking), and (2) determine whether the thinking is relatively favorable or unfavorable (bias in thinking).

Phase 2: Central and Peripheral Routes to Persuasion

Variables such as distraction and personal relevance can determine where a person falls [p. 232 ↓] along the elaboration continuum. At the high end of the continuum, people assess object-relevant information in relation to knowledge that they already possess, and arrive at a reasoned (though not necessarily unbiased) attitude that is well articulated and bolstered by supporting information (the “central route” to judgment where the focus is on assessing information central to the merits of the attitude object). When people are thinking intently, whether the thoughts are favorable or unfavorable are the key determinants of influence and there are many factors that can motivate or enable favorable or unfavorable thoughts. At the low end of the elaboration continuum, information scrutiny is reduced. Nevertheless, attitude change can still result from a low-effort scrutiny of the information available (e.g., examining less information than when elaboration is high or examining the same information less carefully).

Furthermore, if people are generating few thoughts relevant to the merits of the issue, the ELM holds that there are additional change mechanisms that can come into play to influence attitudes. These mechanisms require relatively little in the way of cognitive resources and include processes such as classical conditioning (Staats and Staats, 1958), self-perception (Bem, 1972), and the use of heuristics (Chaiken, 1987). In one early demonstration of different persuasion mechanisms under high and low thinking conditions, Petty et al. (1981) manipulated the personal relevance of a message along with argument quality just as in the study mentioned earlier (Petty and Cacioppo, 1979b). In addition, however, the expertise of the message source was varied (i.e., whether the message on an educational issue was said to come from a Princeton University Professor or a local high school student). Under high relevance conditions, attitudes were determined by the quality of the arguments, just as in the earlier study. The expertise of the source mattered little when thinking was very high. The new result was what happened under low relevance conditions. Here, argument quality made

little difference and attitudes were only affected by source expertise with more attitude change to the high than the low expert source. This study suggested that attitudes were determined by a high thought process – evaluation of the arguments presented – when motivation to think was high, but by a low thought process – reliance on an expertise cue – when motivation to think was low.

In a critique of the ELM, the HSM, and other dual process theories more generally, Kruglanski and Thompson (1999) correctly noted that many of the early studies on dual processes of persuasion (such as the study just described) compared the impact of relatively simple cues (e.g., expertise) described briefly with more complex verbal arguments (e.g., nine consequences of adopting a recommendation each presented in a separate paragraph). This fact led them to suggest that perhaps there was only one mechanism of persuasion that operated and it only appeared as if there were two separate processes because two separate kinds of content were available to process. The problem, as they saw it, was that evidence for dual processes came from studies in which the central route (or high effort processing) resulted from the impact of complex message factors, and the peripheral route (or low effort processing) resulted from the impact of simple source and other nonmessage factors such as one's mood.

However, it is not the case that all dual process studies suffer from this confound. At the conceptual level, in the ELM, content (e.g., source versus message variables; simple versus complex presentation) and process (e.g., effortful scrutiny, reliance on cues) are orthogonal. That is, one can engage in effortful scrutiny for merit of message and source factors, and these features of the persuasion context can also serve as simple cues to persuasion if thinking is low. Similarly, one can process simple or complexly presented material with relatively high or low amounts of effort. Thus, although some ELM research has manipulated simple source [p. 233 ↓] versus complex message variables to study high versus low effort attitude change as Kruglanski and Thompson noted (e.g., Petty et al., 1981), other ELM research has presented only complex message information to show how it could be processed differently depending on whether motivation to think was relatively high or low (e.g., Petty and Cacioppo, 1984b). Furthermore, some research has manipulated very simple to process source factors (e.g., attractiveness) and pointed to its evaluation as an argument under high thinking conditions but as a peripheral cue under low thinking conditions (Petty and Cacioppo, 1984a).

The point is that when a person's goal in scrutinizing all of the information is to determine the true merits of the proposal, the person will use whatever information seems useful in reaching that goal. Thus, if providing a message recipient with extensive information about the source convinces the person more of the validity of the position when the source information is scrutinized, the impact of the source information could be even larger under high than under low thinking conditions. Conversely, if the source information proves irrelevant to the merits of the attitude object when scrutinized (e.g., an attractive source arguing for a new tax law), then its impact will be reduced under high thinking conditions. Illustrating that different processes can be applied to the *same* information under low and high thinking conditions requires a study in which the information serving as a cue and as a substantive argument is exactly the same (i.e., there are no differences in length, complexity of information, placement, etc.). To demonstrate dual processes, all that should vary is how individuals process and use the *same* information that is presented.

In the relevant conditions of one study, Petty and Cacioppo (1984b) compared how people would respond to a message with three strong arguments versus one with three strong plus three weak arguments. If people are carefully processing the arguments, there should be no more persuasion when three weak reasons are added to three strong ones. Indeed, the extent of persuasion could even be reduced as negative issue-relevant thoughts are combined with positive thoughts. If people are not processing the messages carefully, however, then evaluation might occur by a different, less effortful process. People might simply *count* the arguments and reason that six arguments are better than three, leading to more persuasion.

To examine whether the same information could be processed differently leading to different persuasion outcomes under conditions fostering relatively high versus low motivation to think, Petty and Cacioppo (1984b) varied the personal relevance of the message topic along with the message type. When relevance was high, adding weak arguments to strong ones did not enhance persuasion but when relevance was low, adding weak arguments to strong ones led to a significant increase in agreement. This study reveals that even though high and low self-relevance individuals were exposed to the exact same information (three strong plus three weak arguments versus three strong only), they used a different evaluation strategy (i.e., processed the information differently) under high and low relevance conditions leading to very different persuasion

outcomes. Research such as this demonstrates that the same information can be processed in qualitatively different ways depending on a person's overall motivation and ability to think (see Petty and Briñol, 2006; and Petty et al., 1999, for additional discussion of multi-versus single-process models of persuasion).

Phase 3: Elaboration Affects Attitude Strength

According to the ELM, attitudes that are changed with relatively high versus low amounts of issue-relevant thought are postulated to be stronger than attitudes that are changed to the same extent as a result of minimal object-relevant thought. By stronger, [p. 234 ↓] we mean that the attitudes are more likely to persist over time, resist change, and have an impact on other judgments and behavior (see Krosnick and Petty, 1995). This is true regardless of whether the enhanced thinking taking place is relatively objective or biased. There are several reasons for this. First, as thinking increases during attitude change, people should acquire more support for their attitudes (knowledge) and their attitudes should become more accessible and internally consistent. Furthermore, as a result of thinking, people should become more confident in their views. Each of these factors would increase the likelihood that attitudes would be consequential (see Petty and Krosnick, 1995).

The available evidence supports the idea the elaboration enhances attitude strength. For example, in one set of studies, individuals who engaged in greater thinking during attitude formation showed greater persistence over time and more resistance to change when their newly formed attitudes were challenged immediately compared to individuals who formed similar initial attitudes but with less thinking (Haugtvedt and Petty, 1992). However, it is important to note that persistence over time and resistance to change can be independent such as when multiple pairings of an attitude object with positive cues lead it to persist over time, but do not help it resist attack (Haugtvedt et al., 1994). This is because pairing an attitude object with positive cues can make the favorable attitude memorable, but these cues will not help the attitude resist an attacking message that relies on argumentation (see Wegener et al., 2004, for a review).

Once a person's attitude has changed, behavior change requires that the person's new attitude rather than the old attitude or previous habits guide action. If a new attitude is based on high thought, it is likely to be highly accessible and comes to mind automatically in the presence of the attitude object. Therefore, it will be available to guide behavior even if people do not think much before acting (Fazio, 1990). However, even if people do engage in some thought prior to action, research suggests that attitudes based on high thinking are still more likely to guide behavior because these attitudes are held with more certainty and people are more willing to act on attitudes in which they have confidence. So strong is the inferential link between thinking and confidence that people do not have to actually engage in more thinking to attain confidence – they only have to believe they have engaged in more thinking (see Barden and Petty, 2008).

Phase 4: Multiple Roles for Persuasion Variables

We have now seen that there is a continuum of thinking that underlies persuasion and that attitudes can be changed by both high and low thought processes with the former attitudes tending to be more consequential than the latter. In outlining these ideas, we have already highlighted several of the roles that a variable can play in producing persuasion. We have seen that variables can serve as cues or as arguments, or they can affect the extent (amount) or direction (bias) in thinking. A fifth role that variables can play when thinking is high is affecting what people think about their thoughts. Since this is the most recent role for variables that has been studied, we discuss it in more depth in a subsequent section (phase 6). But first, it is important to illustrate the ELM principle that any one variable can affect attitudes in multiple ways.

In describing the roles for variables so far, we have mostly used different variables to illustrate each role. Thus, we have seen how distraction can affect the amount of thinking or that source expertise can serve as a simple cue. However, the ELM holds that any *one* variable can serve in each of these roles depending on a number of other factors. In fact, earlier in this chapter we briefly described how an attractive source or a person's good mood could affect attitudes by different processes in different situations.

[p. 235 ↓] Empirical research has supported this “multiple roles” view. For example, in one study (Petty et al., 1993) placing an advertisement for a pen in the context of a comedy or bland documentary affected attitudes differently depending on whether people were motivated or not to think about the ad. When motivation to think was high, the pleasant feelings from the positive program led people to have more positive thoughts about the product and these thoughts led to more favorable attitudes. When motivation to think was low, however, the good feelings from the program induced more favorable attitudes toward the product without enhancing the favorability of the thoughts generated (i.e., good feelings served as a simple cue). The low thinking results are what would be expected from relatively low effort theories of attitude change such as classical conditioning (Staats and Staats, 1958) or the use of an ‘affect heuristic’ (Chaiken, 1987; Slovic et al., 2002). Under high thinking conditions, however, the indirect influence observed is what would be expected from relatively high effort theories of the use of affect such as the “affect infusion” hypothesis (Forgas, 1995) in which emotions can make retrieval and generation of affectively congruent cognitive material more likely (see Petty et al., 2003, for a review of emotions and persuasion).

According to the ELM, however, these are just two of the roles that variables can play in persuasion settings. When thinking is high, not only should emotions bias the thoughts that come to mind, but also the emotion itself can be evaluated as an argument. The “mood as input” model of emotions was designed to account for just such situations where people scrutinize their emotions as evidence (see Martin, 2000). There is one more process by which emotions can operate when thinking is high – affecting confidence in thoughts (Briñol et al., 2007), and we discuss this role in the sixth phase of ELM research.

Finally, when the likelihood of thinking is not constrained to be high or low by other variables, emotions can affect the extent of thinking. The “mood as information” theory of emotions is one of several theories that makes this prediction. The idea is that negative emotions signal that the world is unsafe or problematic and thus information processing is needed. Positive emotions signal the opposite – that the world is safe and thus thinking is not necessary (Schwarz et al., 1991). If sadness, for instance, leads to more thinking than happiness, then people would actually be more persuaded when sad than happy if the message arguments are strong, but less persuaded when sad than happy if the arguments are weak (Bless et al., 1990).

Although different theories of emotion and judgment have developed around each of the specific roles for variables that the ELM holds to be possible, and some theories of emotions have even considered more than one role (e.g., seeForgas, 2005), no other theory incorporates all of these processes. Perhaps more importantly, unlike the specific theories of emotion, the ELM holds that these same fundamental processes can be applied to a host of other variables such as source attractiveness or recipient power that have nothing to do with emotion.

Phase 5: Extending beyond the Persuasion Context

As described earlier, the ELM was originally proposed as a theory of persuasion (attitude change), but Petty and Cacioppo (1986a) noted that the same principles could be applied to virtually any judgment. Over time, the ELM was used as a framework to study a diversity of persuasive messages on all sorts of topics and in a variety of domains (e.g., health communications, consumer advertisements, legal appeals). A pioneer in moving the ELM beyond persuasion studies was Duane Wegener. Petty met Wegener when the latter came to Ohio State for graduate study in the early 1990s. Wegener was notorious for keeping his advisor (Petty) at work late into the evening with “just one more idea” that he wanted to discuss. Following his PhD, Wegener became a faculty member at Yale, [p. 236 ↓] then Purdue, and he ultimately returned to his alma mater as a faculty member in 2010. Although Wegener developed several influential lines of research that did not involve the ELM (e.g., see Wegener and Petty, 1997), an important ELM contribution was to show that the four core ideas of the ELM outlined above have broader applicability than in the traditional attitude change arena. For example, in the domain of stereotyping, Wegener et al. (2006) showed that a person's existing stereotypes can serve in multiple roles when forming attitudes about a particular member of the stereotyped group. Prior research on stereotyping had focused either on how stereotypes can bias information processing (a high effort process; for example, Kunda and Sherman-Williams, 1993) or on how stereotypes can serve as simple heuristics to judgment (a low effort process; for example, Bodenhausen, 1990).

Wegener noted that according to the ELM, however, both roles for stereotypes should be possible depending on the likelihood of thinking.

In one study demonstrating high and low thought roles for stereotypes, Wegener and colleagues (2006) had college students watch a videotape of a child working on some intelligence test questions in which they could observe the answers the child provided. Prior to the videotape, the students learned that that child came from either a high or a low socioeconomic status (SES) background. When not under cognitive load, higher SES led the students to give higher estimates of the child's intelligence and this was mediated by the thoughts listed about the child consistent with the idea that SES could bias processing of the information observed about the child. However, when under cognitive load, the SES information was also associated with greater estimates of intelligence, but this effect was not mediated by thoughts consistent with the use of SES as a heuristic. In a second study, Wegener et al. (2006) showed differential strength consequences for these judgments. That is, the initial impressions of the child that were influenced by the SES stereotype were more resistant to change by subsequent contradictory information when the initial impressions were formed under high rather than low thought conditions.

Although the studies just described did not use a typical persuasion paradigm, they did involve making evaluative judgments about a target's intelligence. Thus, the ELM might reasonably be expected to operate. What if the judgment requested had nothing to do with evaluation? For a second example of the applicability of ELM principles beyond the persuasion context we turn to another series of studies conducted by Wegener and colleagues, this time on numerical anchoring.

The *anchoring effect* occurs when exposure to a seemingly high (versus low) random number influences participants' numeric responses to a question (Tversky and Kahneman, 1974). For example, if participants are asked to write the last four digits of their social security number on a piece of paper before estimating the age George Washington was when he died, those with high SSNs estimate a higher age than those with low SSNs. Some theories of anchoring assume that the effect occurs by a relatively high effort process conceptually similar to biased processing (e.g., see Mussweiler and Strack, 1999). That is, the anchor biases thoughts in an anchor consistent direction. Other theories, however, assume that anchors work by a less cognitively effortful route.

For example, the anchor could provide a simple hint that the answer is large or small (Schwarz, 1994) or prime a general feeling of high or low quantity which is used to infer the answer (Oppenheimer et al., 2008).

As should be clear by now, the ELM suggests that both high and low effort anchoring processes are possible but would operate at different points along the elaboration continuum. To examine this idea, in one study, Blankenship et al. (2008), asked students whether the answer to a particular question (e.g., the age of Neil Armstrong when he walked on the moon) was higher or lower than a presumably randomly generated high or low number. For some participants, during [p. 237 ↓] the anchoring questions (four with high anchors and four with low anchors), they were given a secondary task to perform that would disrupt the anchor from biasing thinking. Other respondents were not distracted during the anchoring task. Finally, all participants responded to the questions both in the initial session and then one week later. At the delayed questioning, no distraction was present. The results of the study revealed that there was a similar anchoring effect initially for both individuals under high and low cognitive load. However, when asked again one week later, the individuals who had presumably used the anchor thoughtfully (low cognitive load) showed greater persistence of the anchoring bias consistent with the idea that when elaboration is involved, it can enhance the strength of any judgment. In another study, the anchoring effect was also shown to be more resistant to counter influence when it was challenged immediately. Thus, the work by Wegener and colleagues shows that the ELM strength postulate appears to hold beyond the prototypical attitude change domain.

Phase 6: A New Role for Variables – Self-Validation

As we have seen, in the original formulation of the ELM, under the central route to persuasion, much attention was paid to the number and the valence of thoughts people generated to a persuasive message. Other aspects of thoughts, though mentioned briefly in original treatments of the ELM, received scant research attention. However, in the past decade a particular aspect of thoughts has proven to be very important – the overall confidence people have in the thoughts that they generate. Thought confidence

is a meta-cognition that refers to a sense of how valid one's thoughts seem. Thought confidence is consequential because the extent of thought confidence affects whether people use their thoughts in forming their judgments. This idea is referred to as the *self-validation hypothesis* (Petty et al., 2002) and is compatible with the lay epistemic notion (Kruglanski, 1990) that people not only generate ideas, but also seek to determine their correctness.

Research on self-validation might not have occurred had Petty not met Pablo Briñol at a two-day conference on “two roads to persuasion” hosted by the University of Salamanca (Spain) in November of 1998. Briñol was a graduate student in social psychology at the Universidad Autónoma de Madrid (UAM) when he decided to attend the conference to learn more about behavioral factors in persuasion, the intended topic of his dissertation. Briñol approached Petty after his talk to ask some questions and ended up serving as translator for Petty for the remaining talks – all given in Spanish. During the session breaks, the pair planned some studies that were aimed at pinning down the mechanism by which the effects of an earlier behavioral manipulation – head nodding (Wells and Petty, 1980) – affected attitudes. When the results of the planned studies subsequently turned out in a surprising way, the self-validation hypothesis was developed.

Specifically, the research on head nodding, which became Briñol's dissertation under the supervision of Petty and Alberto Becerra, showed that head nodding (moving one's head up or down or side to side during exposure to a message) interacted with argument quality to affect attitudes. This interaction result normally would be interpreted as evidence that head nodding affected the extent of thinking about the message, but there was no evidence that this pattern resulted from differences in the number or nature of the thoughts produced. Rather, it appeared that vertical head movements validated the thoughts that people had, magnifying their impact on attitudes. The argument was that nodding (vs. shaking) one's own head served to validate one's own thoughts similar to how other people nodding (vs. shaking) their heads in response to an individual speaking would validate (or invalidate) what the individual was saying via social consensus (Festinger, 1954). When this research was [p. 238 ↓] written for publication, reviewers found the explanation to be a little odd and unconvincing, so Briñol and Petty, along with a new Ohio State graduate student, Zakary Tormala, decided to conduct some more direct tests of the self-validation idea. Following his

dissertation defense, Briñol joined the faculty at UAM and for every year since he has spent each fall as a visiting scholar at Ohio State. During this period, much progress on the self-validation hypothesis was made.

In the first direct test of the self-validation notion (Petty et al., 2002: Study 1), Ohio State students were asked to list their thoughts on the issue of a new campus proposal and then rate the confidence they had in their thoughts as well as their attitudes on the topic. A key result of this study was that not only were attitudes affected by the number and valence of thoughts listed (as many prior studies had shown), but also by thought confidence. People were more likely to use thoughts in forming their attitudes when confidence in those thoughts was high rather than low.⁴

Once it was clear that thought confidence was an important factor in translating thoughts into attitudes, it suggested that influencing thought confidence would be one more way in which variables can impact attitudes. Demonstrated ways to affect thought confidence now include head nodding (Briñol and Petty, 2003) and many other variables. As one additional example, consider the well-studied variable of source credibility. We have already noted several roles that credibility could play in producing persuasion (e.g., serving as a simple cue when thinking is low, biasing the thoughts message recipients have when thinking is high, etc.). It is now clear that under certain conditions, source credibility can also affect thought confidence.

In one study (Tormala et al., 2006), information about source credibility was presented *after* participants had processed a message containing either strong or weak arguments. The key idea was that people would reason that if the information presented by the source was valid (or invalid as inferred from source credibility), their own thoughts in response to the message would also be valid (or invalid). Consistent with this notion, when the message presented strong arguments and thoughts were mostly favorable, increased source credibility was associated with more persuasion because people relied on their positive thoughts. However, when the message presented weak arguments and thoughts were mostly unfavorable, increased source credibility was associated with less persuasion because people relied on their negative thoughts. In other research examining source credibility effects under high thinking conditions,

source credibility biased thinking when it preceded the message but affected thought confidence when it came after processing was completed (Tormala et al., 2007).

This work suggests that research on persuasion can benefit from considering the timing of the key manipulations as placement of the independent variable in the sequence of persuasion stimuli can have an impact on the mechanism by which it operates. In accord with the ELM multiple roles idea, the self-validation mechanism operates at the high end of the elaboration continuum and occurs when the sense of confidence experienced is most naturally attributed to one's own thoughts, such as when the feeling of confidence is concurrent with or follows thought generation (see Briñol and Petty, 2009, for a review of the many variables that have now been shown to influence thought confidence).

Advantages of the ELM

The ELM is a multi-faceted theory. It points to different attitude change processes that operate in different circumstances. It suggests that any one variable can work in multiple ways and sometimes produce opposite outcomes (e.g., high source credibility leading to more persuasion when it serves as [p. 239 ↓] a cue but to less persuasion when it enhances thinking about weak arguments). It further indicates that the same persuasion outcome can be produced by different processes (e.g., source credibility leading to more persuasion both when it serves as a cue and when it enhances thinking about strong arguments, validates one's favorable thoughts, or biases thoughts). And, it postulates that not all judgmental outcomes that look the same on the surface really are the same (e.g., the same judgments induced by high versus low thinking processes are differentially persistent over time). In the remainder of this chapter we summarize some of the key benefits of such a multifaceted theory for the field of persuasion and beyond.

Coherence in the Field of Persuasion

Integration of Empirical Outcomes

In our view, the ELM has brought some coherence to an attitude change literature that had gotten quite messy. As noted earlier, in the 1970s, numerous scholars complained about the bewildering array of seemingly inconsistent findings in the field and bemoaned the fact that even simple variables could sometimes increase persuasion but at other times reduce it. The ELM explains how and when these different outcomes can occur. It was also confusing that sometimes changed attitudes appeared to be consequential but at other times changed attitudes were not meaningful. The ELM also explains how and when each effect is likely.

In addition to addressing these longstanding puzzles, the ELM has been useful for understanding some current controversies. As one example, consider recent research on implicit measures of attitudes. Contemporary implicit measures aim to assess evaluations that come to mind automatically with little thinking whereas deliberative measures allow some time for reflection (see Petty et al., 2009c). Although the ELM has focused on how the extent of thinking during attitude formation affects whether attitudes are based on central or peripheral processes, it is possible to apply the elaboration continuum idea to the extent of thinking that occurs during attitude expression. Paralleling previous ELM findings, current research is consistent with the idea that simple cues that do not affect attitudes that are reported on deliberative measures often still have an impact on attitudes that are assessed with measures allowing for little thinking (see Petty and Briñol, 2010, for further review).

Integration of Different Theories of Persuasion

Our discussion of the ELM so far has focused on the ELM as a primary theory of judgment. However, the ELM was also intended as a metatheory (theory about theories)

in that it specified what the domain of operation of different theories was. As an early example, Petty and Cacioppo (1986a) noted that the ELM could be used to understand differences between the competing dissonance (Festinger, 1957) and self-perception (Bem, 1972) theories. From the vantage point of the ELM, each of these theories attempted to account for many of the same phenomena (e.g., why people changed their attitudes more when advocating something for a small rather than a large incentive), but did so by very different mechanisms in different situations. Most importantly, self-perception theory relied on a simple inference process and thus it should be more likely to operate on the low end of the elaboration continuum, whereas dissonance theory relied on extensive cognitive activity and thus should be more likely to operate when motivation and ability to think were high. Similarly, we earlier noted how separately developed theories of the impact of emotion on judgment could be organized according to the ELM processes.

Indeed, according to the ELM framework, most of the major theories of attitude change are not necessarily competitive or contradictory, but rather operate in different circumstances. Some theories (e.g., cognitive [p. 240 ↓] response, cognitive dissonance, mood as input) refer to processes that require diligent and effortful information-processing activity, whereas others (e.g., classical conditioning, self-perception, affect heuristic) postulate processes that proceed with considerably less mental effort (see Petty and Cacioppo, 1986b; Petty and Wegener, 1998). The ELM does not diminish the importance of the individual theories. Rather, these theories can be viewed as specifying in more detail the specific process involved under relatively high and low thought conditions. That is, whereas the ELM lumps all kinds of simple cue processes together and all kinds of biased processing theories together, the more specific theories are useful for fleshing out the mechanistic details.

The ELM lumps theories into broad process categories based on the common mechanisms involved, the situations in which they operate, and the consequences observed. For example, cue theories have in common that attitude change moves in the direction of the valence (positive or negative) of the cue, occurs with relatively little thinking, and results in a judgment that is less consequential than a judgment rendered with higher thought. But, the specific way in which this occurs (e.g., conditioned association, use of a heuristic) is also worthy of study. The ELM is designed to be a

general approach that can explain the effects of a wide array of variables that have been examined separately under the rubric of different theories.

Integration of Source, Message, Recipient and Context Variables

Because of the ELM postulate that any one variable can produce persuasion in multiple ways, the classic source, message, recipient, and context variables that affect attitudes can be examined from a common perspective. That is, one can see how very different variables such as source credibility and a person's emotions operate to influence attitudes by the very same fundamental mechanisms. Furthermore, the ELM provides a useful framework for approaching completely novel variables. For example, if one wondered how the color of the paper on which a message was printed would influence attitudes, one would look for simple cue effects when thinking was low (e.g., the most liked color would produce the most favorable attitudes), but would look for other effects (e.g., affecting thinking, biasing thinking, validating thoughts) as the elaboration likelihood was increased.

Furthermore, the ELM can shed new light in looking at traditional variables that the literature appears to have relegated to just one role. Consider the operation of self-relevance. Much research has shown that when the self-relevance of a message is made salient prior to a communication, it influences the amount of thinking (Petty and Cacioppo, 1979b). However, when self-relevance is induced after the message, it affects thought confidence (Petty and Briñol, 2011). Although in this case the two processes lead to a similar result (i.e., a greater argument quality effect under high vs. low self-relevance), the underlying mechanism is quite different.

Real-World Applications of the ELM

A discussion and review of the many areas of application of the ELM is well beyond the scope of this chapter. Thus, we just briefly note that although much ELM research has been conducted in the laboratory, there is considerable work that has been conducted

in field settings as well (e.g., Bakker, 1999). The ELM has proven especially useful in the domains of marketing and advertising (Haugtvedt and Kasmer, 2008; Rucker et al., 2007) and health communication (Briñol and Petty, 2006; Petty et al., 2009a), though there are also applications in the legal, environmental, political, and educational fields as well. Indeed, the ELM has provided practical guidelines for developing effective communications on a wide variety of topics. Tutorials are available to illustrate the actual steps [p. 241 ↓] policymakers and others might take in improving their persuasive appeals using ELM principles (e.g., Briñol and Petty, 2006; Rucker and Petty, 2006).

One of the reasons the ELM has been so widely applied is because persuasion is everywhere, playing an essential role in politics, religion, psychotherapy, education, and day-to-day social interactions. Given that people attempt to persuade others and are also the targets of persuasion, they often wonder about questions such as: are attractive people particularly persuasive? Are experts more persuasive than nonexperts in convincing a jury? Is fear a good emotional tool or is it counterproductive in order to stop people from engaging in risky behaviors? Humans have a longstanding curiosity about such questions and contemporary scholars continue to study these issues as well. The ELM provides answers based on experimental research to many of these questions or suggests ways to initiate new investigations.

We have already noted several of the benefits of focusing on the basic processes underlying effective persuasion. First, identifying the processes by which variables impact attitudes is essential for determining which outcome (increased or decreased persuasion) will occur. Second, we have seen that the process by which an attitude is formed or changed has considerable consequences for the strength of the attitude. Even though both high and low effort processes can sometimes result in the same extent of influence, the attitudes induced by low thinking mechanisms tend to be less stable and predictive of behavior than the ones produced by higher thinking mechanisms. Thus, understanding process is important because it informs us about both immediate and long-term consequences.

As a final illustration of this point, consider our recent research examining whether the principles of the ELM can be applied to the reduction of prejudiced attitudes. Consistent with the ELM, Martin et al. (2011) found that changing attitudes toward stigmatized

groups can be affected by both simple processes that require little thinking and also by traditional elaborative forms of persuasion. Importantly, even when the obtained attitude change was equivalent for processes requiring a low versus a high degree of thinking, there were important benefits of high elaboration prejudice reduction. That is, although both high and low thinking processes were associated with a reduction in the extremity of prejudiced attitudes, the reductions in prejudice produced by high thinking processes were more persistent and resistant to subsequent attacks than equivalent changes produced by less thoughtful mechanisms. As illustrated by this example, the ELM can serve as a basis for, and shed light on, a variety of phenomenon not only relevant to attitude change but also to numerous other judgments, ranging from reducing prejudice to the operation of various heuristics and biases that influence choice and decision making.

Notes

1 The term “elaboration” is used in the theory to connote that people thoughtfully add something to the information provided externally rather than simply mentally rehearsing the original information. In this sense, the term is more restrictive than “cognitive response” (Greenwald, 1968) which would include the former as well as mere restatements of the message.

2 The arguments are developed in pretesting so that strong arguments elicit primarily favorable thoughts when people are instructed to think about them but weak arguments elicit primarily unfavorable thoughts with the same instructions. All arguments are presented as supporting the advocacy but the strong arguments do so in a more compelling way (e.g., pointing to consequences that are more desirable and likely if the advocacy is adopted; see Petty and Cacioppo, 1986a, for an extended discussion).

3 One exception to this is when people feel certain in an ambivalent attitude. In this case, people engage in greater information processing than if they are uncertain of the ambivalent attitude (Tormala et al., 2008). Similarly, if people feel certainty in a doubted attitude, they could engage in [p. 242 ↓] greater information processing than if they felt uncertainty in a doubted attitude (see Wichman et al., 2010).

4 Thought confidence also predicted attitudes above and beyond other aspects of the thoughts listed such as the likelihood and desirability of the consequences inherent in the thoughts.

References

- Allport, G.W. (1935) Attitudes . In C. Murchison (ed.) *Handbook of Social Psychology* , 798–884. Worcester, MA: Clark University Press.
- Bakker, A.B. (1999) Persuasive communication about AIDS prevention: Need for cognition determines the impact of message format . *AIDS Education and Prevention* , 11, 150–162.
- Barden, J. and Petty, R.E. (2008) The mere perception of elaboration creates attitude certainty: Exploring the thoughtfulness heuristic . *Journal of Personality and Social Psychology* , 95, 489–509. <http://dx.doi.org/10.1037/a0012559>
- Baron, R.S., Baron, P.H. and Miller, N. (1973) The relation between distraction and persuasion . *Psychological Bulletin* , 80, 310–323. <http://dx.doi.org/10.1037/h0034950>
- Bem, D.J. (1972) Self-perception theory . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 6, 1–62. New York: Academic Press.
- Blankenship, K.L., Wegener, D.T., Petty, R.E., Detweiler-Bedell, B. and Macy, C.L. (2008) Elaboration and consequences of anchored estimates: An attitudinal perspective on numerical anchoring . *Journal of Experimental Social Psychology* , 44, 1465–1476. <http://dx.doi.org/10.1016/j.jesp.2008.07.005>
- Bless, H., Bohner, G., Schwarz, N. and Strack, F. (1990) Mood and persuasion: A cognitive response analysis . *Personality and Social Psychology Bulletin* , 16, 331–345. <http://dx.doi.org/10.1177/0146167290162013>
- Bodenhausen, G.V. (1990) Stereotypes as judgmental heuristics: Evidence of circadian variation in discrimination . *Psychological Science* , 1, 319–322. <http://dx.doi.org/10.1111/j.1467-9280.1990.tb00226.x>

Brehm, J.W. (1966) A Theory of Psychological Reactance . New York: Academic Press.

Briñol, P. and Petty, R.E. (2003) Overt head movements and persuasion: A self-validation analysis . *Journal of Personality and Social Psychology* , 84, 1123–1139.

Briñol, P. and Petty, R.E. (2005) Individual differences in persuasion . In D. Albarracín, ed. , B.T. Johnson, ed. and M.P. Zanna (eds), *Handbook of Attitudes and Attitude Change* , pp. 575–616. Hillsdale, NJ: Erlbaum.

Briñol, P. and Petty, R.E. (2006) Fundamental processes leading to attitude change: Implications for cancer prevention communications . *Journal of Communication* , 56, 81–104.

Briñol, P. and Petty, R.E. (2009) Persuasion: Insights from the self-validation hypothesis . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 41, 69–118. New York: Elsevier.

Briñol, P. and Petty, R.E. (in press) The history of attitudes and persuasion research. In A. Kruglanski, ed. and W. Stroebe (eds), *Handbook of the History of Social Psychology* . New York: Psychology Press.

Briñol, P., Petty, R.E. and Barden, J. (2007) Happiness versus sadness as a determinant of thought confidence in persuasion: A self-validation analysis . *Journal of Personality and Social Psychology* , 93, 711–727.

Briñol, P., Petty, R.E. and Wheeler, S.C. (2006) Discrepancies between explicit and implicit self-concepts: Consequences for information processing . *Journal of Personality and Social Psychology* , 91, 154–170.

Cacioppo, J.T. and Petty, R.E. (1982) The need for cognition . *Journal of Personality and Social Psychology* , 42, 116–131. <http://dx.doi.org/10.1037/0022-3514.42.1.116>

Cacioppo, J.T., Petty, R.E., Feinstein, J. and Jarvis, W.B.G. (1996) Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition . *Psychological Bulletin* , 119, 197–253. <http://dx.doi.org/10.1037/0033-2909.119.2.197>

- Chaiken, S. (1978) The use of source versus message cues in persuasion: An information processing analysis . Unpublished PhD dissertation, University of Massachusetts.
- Chaiken, S. (1987) The heuristic model of persuasion . In M.P. Zanna, ed. , J. Olson, ed. and C.P. Herman (eds), Social influence: The Ontario Symposium , 5, 3–39. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Chaiken, S., Liberman, A. and Eagly, A.H. (1989) Heuristic and systematic processing within and beyond the persuasion context . In J.S. Uleman, ed. and J.A. Bargh (eds), Unintended Thought , pp. 212–252. New York: Guilford Press.
- Chaiken, S., ed. and Trope, Y. (eds.) (1999) Dual Process Theories in Social Psychology New York: Guilford Press.
- Deutsch, R. and Strack, F. (2006) Duality models in social psychology: From dual processes to interacting systems . Psychological Inquiry , 17, 166–172. http://dx.doi.org/10.1207/s15327965pli1703_2
- Fazio, R.H. (1990) Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework . In M. Zanna (ed.) Advances in Experimental Social Psychology , 23, 75–109. San Diego, CA: Academic Press.
- Festinger, L. (1954) A theory of social comparison processes . Human Relations , 7, 117–140. <http://dx.doi.org/10.1177/001872675400700202>
- Festinger, L. (1957) A Theory of Cognitive Dissonance . Stanford, CA: Stanford University Press.
- Fishbein, M. and Ajzen, I. (1972) Attitudes and opinions . Annual Review of Psychology , 23, 487–544. <http://dx.doi.org/10.1146/annurev.ps.23.020172.002415>
- Forgas, J.P. (1995) Mood and judgment: The Affect Infusion Model (AIM) . Psychological Bulletin , 117, 39–66. <http://dx.doi.org/10.1037/0033-2909.117.1.39>

Greenwald, A.G. (1968) Cognitive learning, cognitive response to persuasion, and attitude change . In A. Greenwald, ed. , T. Brock, ed. and T. Ostrom (eds), Psychological Foundations of Attitudes , pp. 148–170. New York: Academic Press.

Haugtvedt, C.P. and Kasmer, J. A. (2008) Attitude change and persuasion . In C.P. Haugtvedt, ed. , P.M. Herr, ed. and F.R. Kardes (eds), Handbook of Consumer Psychology , pp. 419–436. New York: Erlbaum.

Haugtvedt, C.P. and Petty, R.E. (1992) Personality and persuasion: Need for cognition moderates the persistence and resistance of attitude changes . Journal of Personality and Social Psychology , 63, 308–319. <http://dx.doi.org/10.1037/0022-3514.63.2.308>

Haugtvedt, C.P., Schumann, D.W., Schneier, W. and Warren, W. (1994) Advertising repetition and variation strategies: Implications for understanding attitude strength . Journal of Consumer Research , 21, 176–189. <http://dx.doi.org/10.1086/209391>

Heider, F. (1958) The Psychology of Interpersonal Relations . New York: Wiley. <http://dx.doi.org/10.1037/10628-000>

Kiesler, C.A. and Munson, P.A. (1975) Attitudes and opinions . Annual Review of Psychology , 26, 415–456. <http://dx.doi.org/10.1146/annurev.ps.26.020175.002215>

Krosnick, J.A. and Petty, R.E. (1995) Attitude strength: An overview . In R.E. Petty, ed. and J.A. Krosnick (eds), Attitude Strength: Antecedents and Consequences , 1–24. Mahwah, NJ: Erlbaum.

Kruglanski, A.W. (1990) Lay epistemic theory in social-cognitive psychology . Psychological Inquiry , 1, 181–197. http://dx.doi.org/10.1207/s15327965pli0103_1

Kruglanski, A.W. and Thompson, E.P. (1999) Persuasion by a single route: A view from the unimodel . Psychological Inquiry , 10, 83–110. <http://dx.doi.org/10.1207/S15327965PL100201>

Kunda, Z. and Sherman-Williams, B. (1993) Stereotypes and the construal of individuating information . Personality and Social Psychology Bulletin , 19, 90–99. <http://dx.doi.org/10.1177/0146167293191010>

Martin, L.L. (2000) Moods do not convey information: Moods in context do . In J.P. Forgas (ed.), *Feeling and Thinking: The Role of Affect in Social Cognition* , pp. 153–177. Cambridge, MA: Cambridge University Press.

Martin, M.A., Briñol, P., Horcajo, J. and Petty, R.E. (2011) Changing prejudiced attitudes by thinking about persuasive messages: An attitude strength perspective . Unpublished manuscript.

McGuire, W.J. (1968) Personality and attitude change: An information-processing theory . In A.G. Greenwald, ed. , T.C. Brock, ed. and T.M. Ostrom (eds), *Psychological Foundations of Attitudes* , pp. 171–196. New York: Academic Press.

Mussweiler, T. and Strack, F. (1999) Hipótesis/consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model . *Journal of Experimental Social Psychology* , 35, 136–164. <http://dx.doi.org/10.1006/jesp.1998.1364>

Oppenheimer, D.M., LeBoeuf, R.A. and Brewer, N.T. (2008) Anchors aweigh: A demonstration of cross-modality anchoring and magnitude priming . *Cognition* , 106, 13–26. <http://dx.doi.org/10.1016/j.cognition.2006.12.008>

Petty, R.E., Wells, G.L. and Brock, T.C. (1976) Distraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification . *Journal of Personality and Social Psychology* , 34, 874–884. <http://dx.doi.org/10.1037/0022-3514.34.5.874>

Petty, R.E. (1977) A cognitive response analysis of the temporal persistence of attitude changes induced by persuasive communications . Unpublished PhD dissertation, Ohio State University.

Petty, R.E. and Cacioppo, J.T. (1979a) Effects of forewarning of persuasive intent and involvement on cognitive responses and persuasion . *Personality and Social Psychology Bulletin* , 5, 173–176. <http://dx.doi.org/10.1177/014616727900500209>

Petty, R.E. and Cacioppo, J.T. (1979b) Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses .

Journal of Personality and Social Psychology , 37, 1915–1926. <http://dx.doi.org/10.1037/0022-3514.37.10.1915>

Petty, R.E., Cacioppo, J.T. and Goldman, R. (1981) Personal involvement as a determinant of argument-based persuasion . Journal of Personality and Social Psychology , 41, 847–855. <http://dx.doi.org/10.1037/0022-3514.41.5.847>

Petty, R.E. and Cacioppo, J.T. (1981) Attitudes and Persuasion: Classic and Contemporary Approaches . Dubuque, IA: Wm. C. Brown

Petty, R.E., Ostrom, T.M. and Brock, T.C. (1981) Cognitive Responses in Persuasion . Hillsdale, NJ: Erlbaum.

Petty, R.E. and Cacioppo, J.T. (1984a) Source factors and the elaboration likelihood model of persuasion . Advances in Consumer Research , 11, 668–672.

Petty, R.E. and Cacioppo, J.T. (1984b) The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion . Journal of Personality and Social Psychology , 46, 69–81. <http://dx.doi.org/10.1037/0022-3514.46.1.69>

Petty, R.E. and Cacioppo, J.T. (1986a) Communication and Persuasion: Central and Peripheral Routes to Attitude Change . New York: Springer-Verlag. <http://dx.doi.org/10.1007/978-1-4612-4964-1>

Petty, R.E. and Cacioppo, J.T. (1986b) The Elaboration Likelihood Model of persuasion . In L. Berkowitz (ed.), Advances in Experimental Social Psychology , 19, 123–205. New York: Academic Press.

Petty, R.E. and Cacioppo, J.T. (1990) Involvement and persuasion: Tradition versus integration . Psychological Bulletin , 107, 367–374. <http://dx.doi.org/10.1037/0033-2909.107.3.367>

Petty, R.E., Cacioppo, J.T. and Haugvedt, C. (1992) Involvement and persuasion: An appreciative look at the Sherif's contribution to the study of self-relevance and attitude change . In D. Granberg, ed. and G. Sarup (eds), Social Judgment and Intergroup

Relations: Essays in Honor of Muzafer Sherif , pp. 147–174. New York: Springer/Verlag.
http://dx.doi.org/10.1007/978-1-4612-2860-8_7

Petty, R.E., Schumann, D.W., Richman, S.A. and Strathman, A.J. (1993) Positive mood and persuasion: Different roles for affect under high and low elaboration conditions . *Journal of Personality and Social Psychology* , 64, 5–20. <http://dx.doi.org/10.1037/0022-3514.64.1.5>

Petty, R.E., Haugvedt, C. and Smith, S.M. (1995) Elaboration as a determinant of attitude strength: Creating attitudes that are persistent, resistant, and predictive of behavior . In R.E. Petty, ed. and J.A. Krosnick (eds), *Attitude Strength: Antecedents and Consequences* , pp. 93–130. Mahwah, NJ: Erlbaum.

Petty, R.E., ed. and Krosnick, J.A., ed. (1995) (eds) *Attitude Strength: Antecedents and Consequences* . Mahwah, NJ: Erlbaum Associates.

Petty, R.E. (1997) The evolution of theory and research in social psychology: From single to multiple effect and process models . In C. McGarty, ed. and S.A. Haslam (eds), *The Message of Social Psychology: Perspectives on Mind in Society* , pp. 268–290. Oxford, UK: Blackwell Publishers.

Petty, R.E. and Wegener, D.T. (1998) Attitude change: Multiple roles for persuasion variables . In D. Gilbert, ed. , S. Fiske, ed. and G. Lindzey (eds), *The Handbook of Social Psychology*, Vol . 1 (4th edition), pp. 323–390. New York: McGraw-Hill.

Petty, R.E. and Wegener, D.T. (1999) The Elaboration Likelihood Model: Current status and controversies . In S. Chaiken, ed. and Y. Trope (eds), *Dual Process Theories in Social Psychology* , pp. 41–72. New York: Guilford Press.

Petty, R.E., Wheeler, S.C. and Bizer, G.Y. (1999) Is there one persuasion process or more? Lumping versus splitting in attitude change theories . *Psychological Inquiry* , 10, 156–163. <http://dx.doi.org/10.1207/S15327965PL100211>

Petty, R.E., Briñol, P. and Tormala, Z.L (2002) Thought confidence as a determinant of persuasion: The self-validation hypothesis . *Journal of Personality and Social Psychology* , 82, 722–741. <http://dx.doi.org/10.1037/0022-3514.82.5.722>

Petty, R.E., Fabrigar, L.R. and Wegener, D.T., (2003) Emotional factors in attitudes and persuasion . In R.J. Davidson, ed. , K.R. Scherer, ed. and H.H. Goldsmith (eds), Handbook of Affective Sciences , pp. 752–772. Oxford, UK: Oxford University Press.

Petty, R.E. and Briñol, P. (2006) Understanding social judgment: Multiple systems and processes . Psychological Inquiry , 17, 217–223.

Petty, R.E., Barden, J. and Wheeler, S.C. (2009a) The elaboration likelihood model of persuasion: Developing health promotions to produce sustained behavior change . In R.J. DiClemente, ed. , R.A. Crosby, ed. and M. Kegler (eds), Emerging Theories in Health Promotion Practice and Research (2nd edition), pp. 185–214. San Francisco, CA: Jossey-Bass.

Petty, R.E., Briñol, P., Loersch, C. and McCaslin, M.J. (2009b) The need for cognition . In M.R. Leary, ed. and R.H. Hoyle (eds), Handbook of Individual Differences in Social Behavior , pp. 318–329. New York: Guilford Press.

Petty, R.E., ed. , Fazio, R.H., ed. and Brinol, P. (eds) (2009c) Attitudes: Insights From the New Implicit Measures . New York: Psychology Press.

Petty, R.E. and Briñol, P. (2010) Attitude structure and change: Implications for implicit measures . In B. Gawronski, ed. and B.K. Payne (eds), Handbook of Implicit Social Cognition: Measurement, Theory, and Applications , pp. 335–352. New York: Guilford Press.

Petty, R.E. and Briñol, P. (2011) Effects of self-relevance on persuasion: A self-validation analysis . Unpublished manuscript.

Rucker, D.D. and Petty, R.E. (2006) Increasing the effectiveness of communications to consumers: Recommendations based on the Elaboration Likelihood and attitude certainty perspectives . Journal of Public Policy and Marketing , 25, 39–52. <http://dx.doi.org/10.1509/jppm.25.1.39>

Rucker, D.D., Petty, R.E. and Priester, J.R. (2007) Understanding advertising effectiveness from a psychological perspective: The importance of attitudes and attitude

strength . In G.J. Tellis, ed. and T. Ambler (eds), *The Handbook of Advertising* , pp. 73–88. Thousand Oaks, CA: Sage. <http://dx.doi.org/10.4135/9781848607897>

Schwarz, N. (1994) Judgments in a social context: Biases, shortcomings, and the logic of conversation . In M. Zanna (ed.) *Advances in Experimental Social Psychology* , 26, 125–162. San Diego, CA: Academic Press.

Schwarz, N., Bless, H. and Bohner, G. (1991) Mood and persuasion: Affective status influence the processing of persuasive communications . In M. Zanna (ed.) *Advances in Experimental Social Psychology* , 24, 161–197. San Diego, CA: Academic Press.

Slovic, P., Finucane, M.L., Peters, E. and Macgregor, D.G. (2002) The affect heuristic . In T. Gilovich, ed. , D. Griffin, ed. , and D. Kahneman (eds), *Heuristics and Biases: The Psychology of Intuitive Judgment* , pp. 397–420. New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511808098.025>

Staats, A.W. and Staats, C. (1958) Attitudes established by classical conditioning . *Journal of Abnormal and Social Psychology* , 67, 159–167.

Steele, C.M. (1988) The psychology of self-affirmation: Sustaining the integrity of the self . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 1, 261–302. New York: Academic Press.

Tedeschi, J.T., Schenker, B.R. and Bonoma, T. (1971) Cognitive dissonance: Private ratiocination or public spectacle? *American Psychologist* , 26, 685–695. <http://dx.doi.org/10.1037/h0032110>

Tormala, Z.L., Briñol, P. and Petty, R.E. (2006) When credibility attacks: The reverse impact of source credibility on persuasion . *Journal of Experimental Social Psychology* , 42, 684–691. <http://dx.doi.org/10.1016/j.jesp.2005.10.005>

Tormala, Z.L., Briñol, P. and Petty, R.E. (2007) Multiple roles for source credibility under high elaboration: It's all in the timing . *Social Cognition* , 25, 536–552. <http://dx.doi.org/10.1521/soco.2007.25.4.536>

Tormala, Z.L., Rucker, D.D. and Seger, C.R. (2008) When increased confidence yields increased thought: A confidence-matching hypothesis . *Journal of Experimental Social Psychology* , 44, 141–147. <http://dx.doi.org/10.1016/j.jesp.2006.11.002>

Tversky, A. and Kahneman, D. (1974) Judgment under uncertainty: Heuristics and biases . *Science* , 185, 1124–1130. <http://dx.doi.org/10.1126/science.185.4157.1124>

Wegener, D.T., Clark, J.K. and Petty, R.E. (2006) Not all stereotyping is created equal: Differential consequences of thoughtful versus non-thoughtful stereotyping . *Journal of Personality and Social Psychology* , 90, 42–59. <http://dx.doi.org/10.1037/0022-3514.90.1.42>

Wegener, D.T. and Petty, R.E. (1994) Mood management across affective states: The hedonic contingency hypothesis . *Journal of Personality and Social Psychology* , 66, 1034–1048. <http://dx.doi.org/10.1037/0022-3514.66.6.1034>

Wegener, D.T. and Petty, R.E. (1997) The flexible correction model: The role of naive theories of bias in bias correction . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 29, 141–208. San Diego, CA: Academic Press.

Wegener, D.T., Petty, R.E. and Smith, S. M. (1995) Positive mood can increase or decrease message scrutiny: The hedonic contingency view of mood and message processing . *Journal of Personality and Social Psychology* , 69, 5–15. <http://dx.doi.org/10.1037/0022-3514.69.1.5>

Wegener, D.T., Petty, R.E., Smoak, N.D. and Fabrigar, L.R. (2004) Multiple routes to resisting attitude change . In E.S. Knowles, ed. and J.A. Linn (eds), *Resistance and Persuasion* , pp. 13–38. Mahwah, NJ: Erlbaum.

Wells, G.L. and Petty, R.E. (1980) The effects of overt head movements on persuasion: Compatibility and incompatibility of responses . *Basic and Applied Social Psychology* , 1, 219–230. http://dx.doi.org/10.1207/s15324834bas0103_2

Wichman, A.L., Briñol, P., Petty, R.E., Rucker, D.D., Tormala, Z.L. and Weary, G. (2010) Doubting one's doubt: A formula for confidence? *Journal of Experimental Social Psychology* , 46, 350–355. <http://dx.doi.org/10.1016/j.jesp.2009.10.012>

Wicker, A.W. (1969) Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects . Journal of Social Issues , 25, 41–78. <http://dx.doi.org/10.1111/j.1540-4560.1969.tb00619.x>

<http://dx.doi.org/10.4135/9781446249215.n12>

Handbook of Theories of Social Psychology: Volume 1

A Theory of Heuristic and Systematic Information Processing

Contributors: Shelly Chaiken & Alison Ledgerwood

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "A Theory of Heuristic and Systematic Information Processing"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n13>

Print pages: 246-267

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n13>

[p. 246 ↓]

Chapter 12: A Theory of Heuristic and Systematic Information Processing

Shelly Chaiken, ed. and Alison Ledgerwood, ed.

Abstract

The heuristic-systematic model proposes two distinct modes of thinking about information. Systematic processing involves attempts to thoroughly understand any available information through careful attention, deep thinking, and intensive reasoning, whereas heuristic processing involves focusing on salient and easily comprehended cues that activate well-learned judgmental shortcuts. Heuristic processing is a more efficient and relatively automatic mode of processing, but more often than not confers less judgmental confidence. Systematic processing confers more confidence but is relatively effortful and time-consuming. Thus, individuals tend to engage in heuristic processing unless they are both motivated and able to think carefully about information, in which case the two modes of processing can have additive, attenuating, or interactive effects. Furthermore, both modes of processing can be relatively open-minded, driven by accuracy concerns, or relatively biased, driven by defense or impression concerns. This chapter situates the heuristic-systematic model within its intellectual and personal history, and highlights key empirical findings that support the model's central tenets.

Introduction

Attitudes have been a primary focus of theory and research in social psychology since the 1920s. Nine decades of research have produced a sizeable and complex body of literature that speaks to questions of how people's attitudes are formed, maintained, and changed, and provides an ever-growing foundation upon which new questions arise and new answers unfold. In 1980, a foundation of process-oriented models that explained attitude change based on how people understand and evaluate persuasive argumentation set the stage for one question in particular: Was careful argument scrutiny the only kind of process by which attitude change could occur? Or might we sometimes change our minds in more efficient, but less effortful, ways?

The heuristic-systematic model of persuasion (Chaiken, 1980, 1987; Chaiken et al., 1989, 1996; Chen and Chaiken, 1999) [p. 247 ↓] answers this question by proposing two distinct modes of information processing. The first mode, *systematic processing*, involves attempts to thoroughly understand any and all available information through careful attention, deep thinking, and intensive reasoning (e.g., thinking carefully about the arguments presented, the person arguing, and the causes of the person's behavior). This information is combined and used to guide subsequent attitudes, judgments, and behaviors. For instance, a systematic approach to thinking about a proposed economic policy might involve reading as many magazine and newspaper reports as possible to learn and develop an opinion about the "best" course of action for the economy. The heuristic-systematic model suggests that such systematic thinking entails a relatively high degree of mental effort, and thus requires that a person (1) *can* devote a certain amount of attention to thinking about the issue, and (2) *wants* to devote this attention. Thus, systematic processing is unlikely to occur unless a person is both *able* and *motivated* to do so.

Heuristic processing is much less demanding in terms of the mental work required and much less dependent on having the ability (e.g., enough knowledge and enough time) to think carefully about information. In fact, heuristic processing can be viewed as relatively automatic because it can occur even when people are not motivated and able to deliberately think about a topic. Heuristic processing involves focusing on easily noticed and easily understood cues, such as a communicator's credentials (e.g., expert

versus nonexpert), the group membership of the communicator (e.g., Democrat or Republican), the number of arguments presented (many or few), or audience reactions (positive or negative). These cues are linked to well-learned, everyday decision rules known as *heuristics*. Like other knowledge structures (e.g., stereotypes), heuristics can vary in their availability and accessibility, as well as in their perceived reliability (i.e., the extent to which a particular person perceives a heuristic to be a valid guide for judgment in a given situation; see Chen and Chaiken, 1999; Darke et al., 1998). Moreover, they can be used self-consciously or non-self-consciously: People may consciously decide to invoke a heuristic in order to inform a subsequent judgment, but heuristics can also influence judgments without intention or self-awareness.

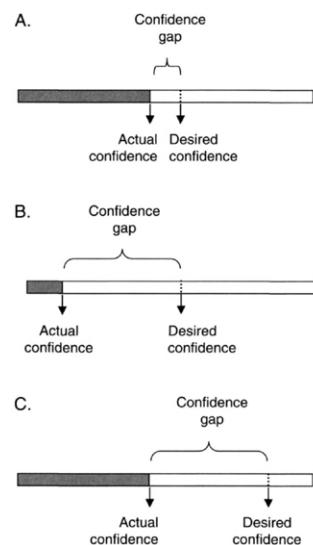
Examples of heuristics include “experts know best,” “my own group can be trusted,” “argument length equals argument strength,” and “consensus implies correctness.” These simple, intuitive rules allow people to form judgments, attitudes, and intentions quickly and efficiently, simply on the basis of the easily noticed cues, and with little critical thinking. A heuristic approach to a proposed economic plan might involve simply adopting the opinion of a noted economist. In other words, heuristic thinking is what we do when we do not have much ability or time to think about something and want to make a reasonable decision as quickly as possible.

The theory further proposed that two principles act in conjunction to determine the mode and extent of information processing that occurs in any given context (Chaiken, 1980, 1987; Chaiken et al., 1989). The model's *least effort principle* reflects the assumption that individuals try to arrive at attitudinal decisions as efficiently as possible (see also Allport, 1954). Thus, all else equal, people should tend to prefer a less effortful mode of processing (i.e., heuristic processing) to one that requires more time and cognitive resources (i.e., systematic processing).

Meanwhile, however, the *sufficiency principle* asserts that individuals are sometimes motivated to exert additional cognitive effort in order to reach a certain level of judgmental confidence. They must therefore balance their preference for maximizing cognitive efficiency with the desire to satisfy their motivational concerns, such as the goal to reach an accurate conclusion (Chaiken et al., 1989; see also Simon, 1976). The heuristic-systematic model suggests that this balance point is determined by a *sufficiency threshold*, defined as the degree of confidence to which an individual aspires

in a given judgmental situation (Chaiken et al., 1989; Eagly and Chaiken, 1993). The sufficiency threshold can be conceptualized as a point located on a continuum of judgmental confidence. The extent of information processing is determined by the size of the discrepancy that exists between an individual's actual level of confidence in their judgment and the sufficiency threshold (i.e., their desired confidence). Thus, effortful information processing should only occur when actual confidence falls below the sufficiency threshold, and should continue (when capacity allows) until this confidence gap is closed. Extent of information processing will therefore depend on both a particular person's actual level of judgmental confidence in a given persuasion setting, as well as their desired level of confidence in that setting (see [Figure 12.1](#)).

Figure 12.1 A person with a small gap between actual and desired confidence might Be able to reach their desired level of confidence (the sufficiency threshold) using only heuristic processing (Panel A). If the confidence gap is larger, either due to a low level of actual confidence (Panel B) or a high level of desired confidence (Panel C), it is less likely that people can reach their desired level of judgmental confidence using only heuristic processing. When people cannot attain their desired level of confidence using only heuristic processing, they will engage in systematic processing in an effort to finish closing the confidence gap, assuming they have the ability to do so



Together, the least effort and sufficiency principles suggest that—assuming adequate cognitive capacity—individuals will engage in systematic processing insofar as the less [p. 249 ↓] effortful heuristic mode does not yield sufficient judgmental confidence (either because heuristic processing cannot occur, as in situations that do not offer easily processed heuristic cues, or because it is insufficient to close the confidence gap). Systematic processing will therefore be increased by factors that either decrease actual confidence, increase desired confidence, or both.

The Multiple-Motive Heuristic-Systematic Model

Although the heuristic-systematic model was initially developed to apply to individuals motivated by accuracy concerns to seek valid judgments, later work expanded the model to incorporate two other broad motivations that can lead to selective information processing geared toward arriving at a particular attitudinal position (Chaiken et al., 1989, 1996; Chen and Chaiken, 1999). The first of these, *defense motivation*, was intended to reflect the impact of such self-focused variables as ego-involvement and personal commitment (see, e.g., Kiesler, 1971; Sherif and Cantril, 1947). According to the multiple-motive model of heuristic-systematic processing, these factors arouse a desire to confirm and defend the validity of preferred attitudinal positions (like one's pre-existing opinion), while challenging the validity of nonpreferred positions. *Impression motivation*, on the other hand, reflects the impact of other-focused variables such as impression-relevant involvement, communication goals, and affiliative concerns (e.g., Higgins and McCann, 1984; Johnson and Eagly, 1989; Smith et al., 1956), which arouse a desire to express attitudes that are socially acceptable.

Like accuracy motivation, defense and impression motivations can engender heuristic and/or systematic processing. However, unlike accuracy motivation, these directional motives tend to lead people to process information selectively, rather than open-mindedly. The biases engendered by these directional motives largely occur outside of awareness; people usually operate under the illusion that they are thinking in an open-minded fashion. In the case of defense-motivated processing, for instance, individuals may selectively choose heuristics that help to confirm a preferred position. A defense-

motivated person might therefore invoke the heuristic “experts know best” if the position of an expert source reinforces her cherished values and social identity, but might choose a different heuristic (e.g., “outgroup sources can’t be trusted”) if the position threatens her social identity. Likewise, impression-motivated heuristic processing entails selective application of heuristics that ensure a smooth interaction with specific others. For example, when interacting with a person or group whose views on an issue are unknown or vague, a perceiver might invoke the heuristic “moderate judgment minimizes disagreement.” On the other hand, when others’ views are known, a “go along to get along” heuristic might better serve the same goal.

With sufficient cognitive capacity and higher levels of motivation, defense-or impression-motivated people will also process systematically, but they will again do so selectively. Thus, a defense-motivated perceiver will attend to, elaborate on, and recall information that serves to bolster his preferred, self-protective position, while an impression-motivated perceiver will systematically process information in such a way as to convey a desired impression to (real or imagined) others.

The multiple-motive model thus proposed three general categories of motives that give rise to three distinct processing goals, any of which can engender heuristic and/or systematic processing. Expanding the theory in this way broadened its applicability to a much wider range of persuasion and social influence situations. In essence, the multiple-motive heuristic-systematic model allowed a rapidly increasing laundry list of persuasion-relevant variables to be understood in terms of their effects on a few key factors – processing goal, cognitive capacity, actual [p. 250 ↓] confidence, and desired confidence—which could in turn suggest a reliable prediction about the extent of attitude change that should occur in a given setting. The strength of this basic dual-process model to organize and generate predictions in the persuasion literature led to its application across a wide range of settings (Chen and Chaiken, 1999; Ledgerwood et al., in press; Mackie, 1987; Stroebe and Diehl, 1988; see Eagly and Chaiken, 1993; Ledgerwood et al., 2006, for reviews). More broadly, it was one of several theories that helped to precipitate a flowering of dual-process models across multiple areas of social psychology, as researchers began to see similar basic principles at work in a number of different domains including stereotyping, impression formation, and decision making (see Chaiken and Trope, 1999).

Personal History of the Theory: An Author by Literature Interaction

In 1972, the first author entered graduate school at the University of Massachusetts Amherst armed with a math major, a psychology minor, and a vague interest in social influence. I began working with Alice Eagly, who was at the time examining the importance of message comprehensibility within the context of Bill McGuire's information processing paradigm. In fact, my masters research project involved testing an idea about comprehensibility that McGuire had tucked away in the depths of a handbook chapter; namely, that the importance of comprehensibility in determining the effectiveness of persuasive appeals would depend on the modality of the communication (Chaiken and Eagly, 1976). Looking back, I can trace part of the development of the heuristic-systematic model to this project. The idea was that message comprehensibility should matter more when the message is in written form rather than audio or video, partly because there is more flexibility to carefully scrutinize a message when reading it than when hearing it spoken. Later, we returned to this idea to examine whether a different type of persuasion variable—source cues—might become increasingly influential as one moved from written to audio and visual modalities (Chaiken and Eagly, 1983). This research suggested that different persuasion variables might be more or less influential depending on how a message was presented.

Here then was one seed for the heuristic-systematic model: different types of persuasion variables had more or less impact depending on a recipient's ability to carefully scrutinize a message. Other seeds were in the recent and current literature at that time: articles on correspondent inference theory, Kelley's covariation theory, and self-perception theory populated the reading lists for my coursework, and I was intrigued both by attribution models and by the simplicity of self-perception (Bem, 1972; Jones and Davis, 1965; Kelley, 1972, 1973). With Alice, I helped develop an attribution model of source characteristics based on Kelley's (1973) cube model, particularly his notions of discounting and augmentation. Yet compared to self-perception theory, analyzing the covariances of even a somewhat superficial persuasion variable like

source characteristics seemed effortful and deliberative. Could people really be so careful, so thoughtful, all the time?

The simple if-then's of self-perception theory appealed to me—why engage in some arduous analysis of your own thoughts and behaviors when you could simply reason: if I'm yelling, I must be angry? Years before, when Kennedy and Nixon had been running for president, I remember listening to my parents consider the intricacies of the various political issues at stake. Meanwhile I (with a young child's preference for the simple that I still haven't seemed to grow out of completely) *knew* that Kennedy was the man to vote for; after all, he looked better. And it wasn't just me; in graduate school, I read about data showing that although those who heard the first Kennedy-Nixon debate on the [p. 251 ↓] radio believed Nixon had won, those who watched it on television were convinced that Kennedy had in fact prevailed. If the average person was really what McGuire (1969) had dubbed "the lazy organism," might a simple if-then suffice for most of us, much of the time?

Then, in 1975, I came across Shelley Taylor's recently published dissertation, which shed some additional light on self-perception processes. Female participants were shown pictures of three different men who varied in attractiveness. Some participants were given false physiological feedback suggesting that they showed a strong preference for one of the men pictured. Participants were also led to believe that they would have the opportunity to meet one of these men in a few weeks (high consequences condition), or were not led to expect a future meeting (no consequences). They then rated each of the three men on attractiveness. The results suggested that participants in the no consequences condition engaged in self-perception: they based their ratings of attractiveness on the physiological feedback provided. Most interestingly to me, however, participants in the high consequences condition were not affected by the feedback manipulation. Instead, there was some evidence to suggest they were thinking more carefully and critically about the three candidates: they spent more time looking at the pictures, and content analyses suggested that they made more critical comments.

I remember thinking to myself that surely this could apply to persuasion. High and low consequences provided a variable that could perhaps predict when a lazy organism would opt for a simple "if-then" versus a more complicated analysis of available

information. I built my main dissertation experiment around this idea, testing whether high versus low consequences would moderate the persuasive impact of source cues (the most frequently studied noncontent variable at the time) and content (extent of strong persuasive argumentation). I reasoned that source cues such as likeability can be processed quite easily and efficiently by a lazy organism unmotivated by future consequences. When future consequences were present, however, participants should be motivated to process information more carefully, and extent of strong argumentation should play a greater role in persuasion.

And (amazingly, to me at the time), the study worked. I started calling the more deliberative mode of thinking systematic, but was unsure what to call the other one until leek Ajzen, another important mentor for me in graduate school, suggested the name “heuristic.” As I continued the line of research, the notion of consequences became abstracted into motivation to process information. Like many other theories at that time, the default motivation was implicitly assumed to be accuracy; I began to explicitly label the motivation “accuracy motivation” only later in order to emphasize that both modes of thinking served the same motive (rather than one being rational and the other irrational). Drawing on my earlier modality research, I also added capacity as a second variable that seemed necessary for deliberative processing to occur.

Over the years, I tried to expand the model to other kinds of cues, and to test its assumptions in various ways. Perhaps inevitably, given that my intellectual genes were steeped in classic functional theories of attitudes (Alice Eagly had worked with Herb Kelman), it occurred to me that accuracy was not the only motive in town, and I began to try to group the major attitudinal functions I had learned about in graduate school into a few broad categories of motives. Over time, we developed and tested predictions deriving from a multiple-motive heuristic-systematic model that included not just accuracy motivation, but also impression and defense motives (see Chaiken et al., 1996; Chen and Chaiken, 1999, for reviews).

Very gradually, then, the theory expanded—first under the influence of functional theories, and then following new developments in social cognition. I had always thought of heuristics as simple decision rules [p. 252 ↓] that were likely represented in memory, and such a conceptualization lent itself to new theorizing about availability and accessibility in social cognition. By now I was at New York University, where

hallway conversations with Tory Higgins and John Bargh inevitably turned toward basic principles of social cognition. I began to think that heuristics ought to vary in their availability, accessibility, and reliability, and that this would have important consequences for when a given heuristic would be applied. Furthermore, heuristics seemed to me to be relatively automatic, in at least some senses of the term (see Bargh, 1994). I always thought of them as a kind of shortcut; thus, at the very least they were automatic in the sense of being efficient. It also seemed likely that they often (but not always) operated outside of awareness. Over time, the results of accumulating studies provided support for this social-cognitive side of the model as well (see Chen and Chaiken, 1999, for a review).

Conceptualizing heuristics as a form of automatic social cognition highlights one way in which the basic processes underlying the heuristic-systematic model extend beyond the persuasion context to other domains. It became apparent early on that a dual-process perspective was not restricted to a persuasion context; that it would be fruitful to look across different domains to understand the common mechanisms at work in all of them. And indeed, the heuristic-systematic model was just one of a growing family of dual-process models that began to populate social psychology in the 1980s and 1990s, as researchers across different domains converged on a similar set of mechanisms to explain information processing in a variety of settings (see Chaiken and Trope, 1999).

Intellectual History of the Theory

Like many of the models in this family, the heuristic-systematic model suggests that individuals can think about information not only via a bottom-up, data-driven process but also via a more top-down process that depends on the pre-existing knowledge structures they bring to a particular context. Although this represented a radical reorientation in the field of persuasion at the time, the notion that we can rely on learned associations to structure understanding emerged as early as 1930 in Kohler's discussion of sensation and perception, in which he suggested that our perceptions are shaped as much by a top-down application of knowledge derived from past experiences as by bottom-up, sensory experience (see also Moskowitz et al., 1999; Yates, 1985). For instance, upon sensing a pattern of colors and lines with our eyes, we can draw on our past experiences and associations with this pattern to label it a "chair" and

infer its form and function. Subsequently, Bruner's "new look" emphasized the notion that our perceptions are substantially shaped by expectation and motivation (Bruner, 1957). Research on mental schemas developed this idea to suggest that we can quickly organize and "fill in the blanks" about our world using generalized mental structures built from our past experiences (e.g., Anderson and Pichert, 1978; Brewer and Treyens, 1981; see Fiske and Linville, 1980; Taylor and Crocker, 1981; and Fiske and Taylor, 2008, for reviews). Together, these literatures highlight a relatively quick, efficient, top-down method of understanding the world that capitalizes on past experience to structure current understanding, and suggests that these mental shortcuts may be applied to a range of different domains (see also Tversky and Kahneman, 1974).

Within this historical context, the heuristic-systematic model proposed that individuals might sometimes rely on quick, efficient, cognitive shortcuts to make judgments about the validity of information they encounter. Thus, rather than carefully scrutinizing any and all available information, people might instead draw on simple if-then associations learned through repeated experience to [p. 253 ↓] inform their attitude judgments. For instance, given that experts tend to be correct, individuals might develop a learned association between experts and correctness that allows them to easily and efficiently infer that a subsequently encountered expert is likely to be right ("if expert, then correct").

As noted earlier, because the model assumed that heuristics are like other knowledge structures, it invited connections to social-cognitive research on the principles governing the activation and use of stored knowledge (Chaiken et al., 1989, Chen and Chaiken, 1999). In other words, heuristics should be subject to the same principles of availability, accessibility, and applicability that underlie the use of stored knowledge in other domains (e.g., Higgins, 1989; Higgins et al., 1982). Considerable research supports this claim (see Chen and Chaiken, 1999, for a review). For instance, in order to be used to inform attitudes in a given setting, a heuristic must be (1) accessible (e.g., because it has been situationally primed), and (2) applicable (e.g., because an individual believes it to be a reliable, or usable, guide for judgment; Chaiken et al., 1992).

Heuristic processing thus represented a very different mode of thinking from the more systematic, comprehensive mode that had occupied the center stage of persuasion theory and research for some time. Furthermore, the heuristic-systematic model

suggested that these modes of processing involved a tradeoff between optimal judgments (maximized by systematic processing) and efficient judgments (maximized by heuristic processing). The model's original formulation proposed that heuristic or systematic processing would predominate depending on the relative importance of accuracy or economic concerns for a given person in a given context (Chaiken, 1980). Subsequently, this notion was refined to emphasize a continuum of judgmental confidence, along which two critical points can be located: a person's actual confidence, and their desired confidence or *sufficiency threshold* (Eagly and Chaiken, 1993). As lazy organisms (McGuire, 1969), people first attempt to close this gap in confidence via heuristic processing. Only when this easier strategy fails to confer sufficient judgmental confidence will people exert the cognitive effort required by systematic processing, assuming they are able to do so.

Considerable research supports this central claim that individuals will process information heuristically unless they are both motivated and able to engage in more effortful systematic processing. Heuristic cues alone tend to guide judgments when ability is low (such as when participants possess little knowledge about the topic, when they are under time pressure, or when situational constraints diminish cognitive capacity) and when motivation is low (such as under conditions of low task importance or personal relevance; Giner-Sorolla et al., 2002; Petty et al., 1976; Ratneshwar and Chaiken, 1991; Wood et al., 1985). As ability and motivation increase, systematic processing plays an increasing role in influencing attitudes (e.g., Chaiken, 1980; Martin et al., 2007; Petty and Cacioppo, 1984; see Eagly and Chaiken, 1993, for a review). Importantly, the processing modes are by no means mutually exclusive: given adequate levels of ability and motivation, heuristic and systematic processing often co-occur (Chaiken, 1980, 1987; Eagly and Chaiken, 1993). We return to this assumption of *concurrent processing* later in the section.

Bridging beyond the Persuasion Context

Although the heuristic-systematic model was initially developed within the context of the paradigmatic persuasion experiment, in which a source conveys a message to a target with some effect, it quickly became clear that the fundamental processes at work within this context were mirrored in other domains. At its heart, the persuasion paradigm

involves individuals making judgments in light of information, as they do in many [p. 254 ↓] other domains. Because it focuses on the basic processes underlying persuasion effects, the heuristic-systematic model provided a natural bridge from persuasion to many other, conceptually similar, areas. Across various domains, individuals can make judgments based on quick shortcuts or more effortful, extensive processing, and motivation and ability play a key role in guiding the extent to which effortful processing occurs.

Indeed, as noted earlier, the heuristic-systematic model was among several early dual-process models in social psychology. Together, these paved the way for a rapid proliferation of information-processing theories in a variety of domains that distinguished between a relatively automatic, fast, reflexive mode of thinking based on well-learned associations, and a more controlled, analytic, effortful mode based on systematic reasoning (e.g., Brewer, 1988; Devine, 1989; Fazio and Towles-Schwen, 1999; Fiske et al., 1999; Gawronski and Bodenhausen, 2006; Gilbert, 1989; see Chaiken and Trope, 1999; Smith and DeCoster, 2000, for reviews). In their 1999 volume, Chaiken and Trope brought together a variety of dual-process models from diverse fields that converged in their basic distinction between these two types of processes, illustrating that these perspectives are really a family of theories with a common core.

Empirical Findings

The first experiment designed to test the heuristic-systematic model examined whether involvement would moderate the extent to which a heuristic cue (communicator likeability) versus message content (extent of supportive argumentation) affected people's attitudes (Chaiken, 1980: Study 1). Undergraduate participants read a transcript of an interview with a university administrator who in the course of the interview either praised undergraduates (likeable source condition) or disparaged them (unlikeable source condition). Later, the administrator stated his opinion on an issue (e.g., changing from a semester to a trimester system) and provided either a weak message (containing only two arguments) or a strong message (containing six different arguments) in support of his opinion.

To test whether participants' level of motivation would determine the extent to which they relied on the heuristic cue or engaged in more effortful processing of message content, the experiment also manipulated participants' involvement by leading them to expect that they would discuss either the same issue or a different issue at a subsequent experimental session. Participants who expected to discuss the same issue should be more motivated to reach an accurate conclusion about whether the administrator's position was valid, compared to those who expected to discuss a different issue, and should therefore engage in more systematic processing. Consistent with the study's hypotheses, high involvement participants showed greater attitude change in response to a strong (versus weak) message, but were unaffected by communicator likeability. In contrast, low involvement participants showed greater attitude change in response to the likeable (versus unlikeable) communicator, but were unaffected by message content. Furthermore, substantiating the notion that attitude change was mediated via systematic processing in the high involvement condition, these participants showed greater recall of arguments and reported more issue-relevant thoughts, compared with those low in involvement. Thus, which factors produced persuasion—and *how* they produced persuasion—depended critically on participants' level of motivation.

Importantly, by delineating the dual processes underlying people's thinking about persuasive appeals, the heuristic-systematic model was able to shed light on the role played by motivational variables, as well as source cues and message content, in influencing attitudes. For instance, previous research had reported conflicting findings [p. 255 ↓] regarding the impact of involvement on persuasion (e.g., Pallak et al., 1972; Sherif and Hovland, 1961). Our results (Chaiken, 1980) suggested that involvement could either increase or decrease attitude change in response to a persuasive message, depending on the valence of available heuristic cues and the strength of the message content. Similarly, Axsom et al. (1987) showed that whereas involvement increased the impact of argument quality on persuasion, it decreased the impact of the heuristic cue of audience response (i.e., whether an overheard message audience sounded enthusiastic or unenthusiastic). The heuristic-systematic model thus provided a theoretical framework within which to organize a large number of persuasion-related factors in a literature that had often produced contradictory results.

The Concurrent Processing Assumption

It was in large part the prevalence of such contradictory results that motivated the development of the heuristic-systematic model. Looking back, the historical assumptions discussed earlier, combined with the current climate in the persuasion literature, created a unique context within which the logic of a dual-process perspective was perhaps more likely to be discovered. And, in fact, two dual-process models of persuasion independently emerged from this context: the heuristic-systematic model and the elaboration-likelihood model (ELM; Petty and Wegener, 1999). Both provided an organizing framework for understanding the impact of various persuasion variables by suggesting two routes to persuasion: the heuristic or “peripheral” route, and the systematic or “central” route. However, they differed in some important ways. For instance, whereas the ELM assumed that the peripheral and central routes to persuasion were mutually exclusive, the heuristic-systematic model suggested that they could co-occur and even interact.

Thus, although many of the initial dual-process studies of persuasion suggested that heuristic cues do not impact attitudes when people are motivated and able to process systematically (e.g., Axsom et al., 1987; Chaiken, 1980; Petty et al., 1981; Wood et al., 1985), the heuristic-systematic model suggested that this pattern was only one possible outcome of the two modes of information processing. Specifically, these results seemed to represent cases in which systematic processing *attenuated* the judgmental impact of heuristic processing because it took into account information that contradicted the valence of the available heuristic cues. If systematic processing instead yielded information that was congruent with heuristic processing, the heuristic-systematic model suggested an *additivity hypothesis* whereby heuristic processing could exert a direct effect on judgment over and above the impact of systematic processing. Supporting this hypothesis, Maheswaran and Chaiken (1991; see also Maheswaran et al., 1992) found that when heuristic cues and message content were congruent, attitude change was mediated by both heuristic and systematic processing.

Importantly, however, the heuristic-systematic model proposed that the two processes could not only co-occur, but could also interact to exert interdependent effects on judgment. Specifically, heuristic processing could *bias* systematic processing by

influencing people's expectations about the validity of arguments presented in a persuasive appeal (Chaiken et al., 1989). To test this notion, Chaiken and Maheswaran (1994) presented participants with a novel attitude object (a new telephone answering machine called the "XT-100") and assigned them to one cell of a 2 (accuracy motivation: low versus high) by 2 (heuristic cue: valid versus invalid) by 3 (argument quality: strong versus ambiguous versus weak) design. This study manipulated accuracy motivation by varying the importance and personal relevance of participants' decisions regarding this new product. Whereas participants in the high [p. 256 ↓] importance condition learned that they were part of a small list of respondents, that their input would be heavily weighted, and that the product would be distributed in their geographical area, participants in the low importance condition learned that they were part of a large group of respondents, that individual opinions were unimportant, and that the product would be distributed in a different geographical area.

Participants next received a positive message about the product that contained a heuristic cue conveying either high or low validity. Specifically, they learned that the product description in the message was taken from *Consumer Reports*, a credible source, or from a promotional Kmart pamphlet, a noncredible source. The product description contained either strong arguments, weak arguments, or an ambiguous mixture of the two. Participants then reported their attitudes toward the XT-100 and listed their thoughts about the product description.

As in previous studies (e.g., Chaiken, 1980), the relatively unmotivated participants in the low importance condition expressed attitudes that reflected the source credibility cues, but not the quality of the arguments presented in the product description. Thus, participants were more favorable toward the XT-100 when they had read a positive message from a credible (versus noncredible) source, regardless of actual message content. Moreover, this effect of source cue on attitudes was direct, rather than mediated by cognitive elaboration, consistent with the notion that participants were directly inferring the validity of the message from the source's credibility (i.e., processing heuristically by using a well-learned association between credibility and correctness).

Meanwhile, the results for participants in the high importance condition who read an unambiguous message also replicated past research: highly motivated participants who read a strong (versus weak) persuasive message expressed more positive

attitudes toward the XT-100, and this effect was mediated by participants' cognitive elaborations about the product. Additional analyses revealed that when source cue and message content were contradictory in their implications for message validity (i.e., a credible source paired with weak arguments, or a noncredible source and strong arguments), systematic processing alone determined attitudes. This is consistent with the *attenuation hypothesis* suggesting that systematic processing can override the effects of heuristic processing. However, when source credibility and message content were congruent (i.e., a credible source and strong arguments, or a noncredible source and weak arguments), there was both a direct effect of the heuristic source cue on attitudes and an effect of message content mediated by systematic processing. Thus, when the information provided by heuristic and systematic processing were congruent, the results supported the additivity hypothesis suggesting that both modes of processing can independently influence attitudes.

Finally, highly motivated participants who read an ambiguous message were influenced both by the source cue and by systematic processing of the high (versus low) quality arguments. Supporting the *bias hypothesis*, these participants' cognitive elaborations about the attitude object were influenced by the validity information provided by the source cue, such that the high credibility source biased systematic processing in a positive direction, whereas the low credibility source biased systematic processing in a negative direction. In addition, attitudes in this condition were also directly influenced by the heuristic cue.

In other research examining the bias hypothesis, Darke et al. (1998) studied the impact of consensus information presented in the absence of persuasive argumentation on college students' support for comprehensive exams. Accuracy motivation was manipulated via personal relevance. Participants in the high relevance condition were led to believe that the exam policy would have direct personal consequences (i.e., it would take effect the following academic year, and [p. 257 ↓] would thus apply to current students), whereas those in the low relevance condition were led to believe that there would be no personal consequences (i.e., the policy would take effect in ten years, and therefore have no impact on current students). Participants then learned that 80 percent of students either supported or opposed instituting comprehensive exams, based on either a small poll (a sample size of ten students) or a large poll (a sample size of 1,000 students). Consistent with the bias hypothesis, participants in the high

personal relevance condition generated thoughts that were biased in the direction of the available consensus cue, and these thoughts then influenced their attitudes. In contrast, the consensus information exerted a direct, heuristic influence on participants' attitudes in the low personal relevance condition. Interestingly, highly motivated participants also discriminated between the more and less reliable heuristic cues: participants in the high relevance condition were more persuaded by the consensus information when the poll was based on a large versus small sample of students, whereas participants in the low relevance condition were persuaded by consensus information regardless of the poll's size.

Together, then, these studies highlight the complex interplay between heuristic and systematic processing (see also Chen et al., 1996; Erb et al., 1998; Ziegler et al., 2005). Importantly, they demonstrate that the two modes of processing can influence attitudes both independently and interactively, suggesting that they may best be conceptualized as two interdependent and potentially co-occurring ways of thinking (see Eagly and Chaiken, 1993: Chapter 7, for further discussion).

Multiple Motives

Another unique feature of the heuristic-systematic model is that it jointly considers the influence of multiple modes of processing on the one hand and multiple motives on the other. The tripartite analysis of motives in the heuristic-systematic model has its historical roots in the literature on attitude function, although it should be noted that similar classes of motives that center on understanding, protecting the self, and affiliating with others are echoed across multiple domains (e.g., Allport, 1954; Baumeister and Leary, 1995; Deutsch and Gerard, 1955; Fiske, 2002; Heider, 1958; Lerner and Tetlock, 1999; Tesser and Campbell, 1983). The notion that individuals are often motivated to form and hold attitudes that square with relevant facts built on Katz's (1960) knowledge function and Smith et al.'s (1956) object appraisal function of attitudes, which emphasized the role often played by attitudes in organizing experience and guiding action with respect to an individual's ongoing concerns. The heuristic-systematic model was thus initially designed to apply to persuasion contexts in which the message recipient is concerned with assessing the validity of a persuasive appeal (Chaiken, 1980, 1987; Chaiken et al., 1996). We subsequently extended the model

beyond validity-seeking persuasion contexts, adding impression and defense motives to encapsulate two other broad classes of attitude functions in the literature (Chaiken et al., 1989). The concept of impression motivation was designed to capture other-oriented, affiliative functions such as Smith et al.'s (1956) social adjustment function, which emphasized the role that attitudes can play in helping people establish and maintain relationships with other individuals or groups (see also McGuire, 1969). Meanwhile, defense motivation encapsulated self-oriented defensive functions such as Katz's (1960) ego-defensive function and Smith et al.'s (1956) externalization function, which suggested that some attitudes serve to protect individuals' self-image against internal or external threats.

Considerable evidence supports the notion that impression motivation can guide heuristic and systematic processing (see Chaiken et al., 1996, for a review). For example, Chen [p. 258 ↓] et al. (1996: Study 2) led participants to anticipate a discussion about a social issue with a partner who ostensibly held either a favorable or an unfavorable opinion on the issue. Before this discussion, participants read a series of fictitious scenarios designed to prime either the accuracy goal of determining a valid opinion, or the impression goal of getting along with other people. After this task, participants familiarized themselves with the discussion issue by reading an evaluatively balanced essay concerning the issue (in this case, whether election returns should be broadcast while polls are still open). Participants then listed the thoughts that had occurred to them as they read the essay and indicated their own attitudes toward the issue.

Impression-motivated participants expressed attitudes that were much more congruent with their alleged partners' attitudes than did accuracy-motivated participants: when the partner favored one side of the issue, they favored the same side, whereas when the partner opposed it, they opposed it. Interestingly, accuracy-motivated and impression-motivated participants exhibited the same amount of systematic processing (as measured by the number of issue-relevant thoughts that were listed). However, whereas accuracy-motivated participants' systematic processing was open-minded and unbiased by their partners' attitudes, impression-motivated participants exhibited systematic processing that was biased toward their partners' attitudes. For example, when the partner favored allowing broadcasts of election returns while the polls were still open, impression-motivated participants listed thoughts that revealed much more

favorable thinking about arguments supporting the broadcasting of returns and more unfavorable thinking about arguments opposing it.

Like impression motivation, defense motivation can also guide heuristic and systematic processing in a directional fashion, as individuals attempt to close the gap between actual and desired confidence that a judgment will protect their cherished beliefs and self-views (e.g., Ditto and Lopez, 1992; Giner-Sorolla and Chaiken, 1997; Liberman and Chaiken, 1992; Lord et al., 1979). For instance, Giner-Sorolla and Chaiken (1997) found that participants' vested interest in a campus issue biased their judgments of a consensus cue's reliability, when additional information that would permit systematic processing was unavailable. Specifically, participants rated the consensus information (an opinion poll of their fellow students) as more reliable, and criticized it less, when the poll results supported rather than opposed their vested interests. When additional information was available, participants also displayed a defensive bias in their systematic processing, cognitively elaborating the arguments presented in a selective manner that reflected their vested interests. Interestingly, when both types of information were available, exposure to a hostile consensus cue appeared to undermine judgmental confidence and increase systematic processing of the arguments presented: In these conditions, the influence of vested interests on participants' subsequent attitudes was mediated by their cognitive elaborations about the issue. In contrast, exposure to a congenial cue appeared to close the confidence gap, such that participants simply used their vested interests to directly inform their subsequent attitudes, rather than engaging in additional heuristic or systematic processing. Thus, as with accuracy and impression motives, both heuristic and systematic processing can be used to serve self-protective processing goals.

Implications for Social Issues

Because it focuses on the basic mechanisms by which persuasion can occur, the heuristic-systematic model can predict how a wide range of variables will influence attitudes and judgments in various situations. It is therefore a particularly powerful tool for understanding and influencing information [p. 259 ↓] processing in ways that can help effect positive social change, and has been applied to diverse issues such as increasing individuals' acceptance of potentially threatening health information,

improving the design of product warning labels, identifying and decreasing bias in jury decision-making, increasing recycling behavior, and developing more effective programs for preventing substance abuse among teens (e.g., Brewer and Hupfeld, 2004; ForsterLee et al., 2006; Harris and Napper, 2005; Howard et al., 2006; Jepson and Chaiken, 1990; Liberman and Chaiken, 1992; Scott, 1996; Werner et al., 2002; Zuckerman and Chaiken, 1998). Here, we discuss the implications of the heuristic-systematic model for two areas that we find particularly interesting: negotiation and political decision-making.

Negotiation and Conflict Resolution

Research exploring heuristic and systematic processing in simulated negotiations has confirmed the utility of a dual-process perspective for understanding information processing in conflict settings (see Ledgerwood et al., 2006, for a review). Specifically, when negotiators have modest levels of motivation (or low cognitive capacity), they often rely on heuristics such as fixed-pie assumptions (the perception that a negotiation is a zero-sum game), initial anchor values (e.g., first offers, or information about the typical outcome of similar negotiations), and stereotypes about an opponent's group membership (De Dreu et al., 1999; Thompson and Hastie, 1990; see De Dreu, 2004, for a review). In contrast, when motivation and capacity are relatively high, sole reliance on these heuristics tends to decrease as systematic processing increases.

Researchers have identified several factors that influence the extent to which people process information in negotiations (see De Dreu, 2004). These factors include both stable individual differences and temporary elements of a given situation that influence motivation and/or capacity. For instance, negotiators who are high in the dispositional need for cognitive closure—that is, the desire to reach a judgment quickly and avoid ambiguity (Webster and Kruglanski, 1994)—are more likely to rely solely on heuristics than are those who have a low need for closure (De Dreu et al., 1999).

Temporary, situation-specific factors such as the presence of a highly involving task or process accountability (the need to justify the way in which a decision is made) tend to increase the extent of systematic processing, whereas time pressure and capacity-degrading conditions (e.g., noise) tend to decrease such processing (e.g., De Dreu,

2003; Tetlock et al., 1989; see Ledgerwood et al., 2006, for a review). For example, De Dreu (2003) examined the effect of time pressure on fixed-pie perceptions. Business students were placed into pairs and asked to play the role of a buyer or seller in a negotiation over the purchase of a car. The negotiation task was designed to hold integrative potential: the different issues varied in importance to the two negotiators, so that an integrative solution that capitalized on this variation in priorities would be more beneficial to both negotiators than a 50:50 split based on a fixed-pie assumption. Participants were led to believe that they had either plenty of time to complete the negotiation (low time pressure condition), or relatively little time (high time pressure condition). Participants were more likely to revise their fixed-pie assumptions, which led to higher joint outcomes, under low rather than high time pressure. These results suggest that time pressure reduces systematic processing, heightening reliance on heuristic cues such as fixed-pie perceptions and preventing negotiators from capitalizing on integrative potential.

In contrast, when an individual expects to discuss an issue with, justify a decision to, or be evaluated by an unknown audience, he or she tends to engage in pre-emptive [p. 260 ↓] self-criticism, displaying a heightened motivation to arrive at an accurate conclusion (Lerner and Tetlock, 1999; Tetlock et al., 1989). In terms of the heuristic-systematic model, holding a person accountable to an audience whose views are unknown can increase desired confidence for a correct judgment and thereby stimulate accuracy-motivated systematic processing. Confirming this idea, De Dreu et al. (2000) randomly assigned business student participants to high-accountability and low-accountability conditions before asking them to engage in a mock negotiation over the purchase of a car. In the high-accountability condition, participants expected that their negotiation strategies and decisions would be reviewed and evaluated several days later by an experienced negotiator and a psychologist. In the low-accountability condition, participants did not receive this information. The results showed that under high accountability, participants were more likely to revise their fixed-pie assumptions and tended to obtain higher joint outcomes. Together, these studies suggest that negotiation outcomes can be improved by reducing the impact of variables that decrease accuracy motivation and capacity (like time pressure), as well as by facilitating factors that increase accuracy motivation (like accountability to an impartial expert).

Political Attitudes

The heuristic-systematic model can also be used to shed light on political decision-making and voting behavior (e.g., Forehand et al., 2004; Marcus et al., 2000; Mondak, 1993; Newman and Perloff, 2004), and suggests that the impact of various factors on political judgments and intentions will depend on a voter's ability and motivation to think about available information. When people are motivated and able to process political information, they will tend to weigh the quality of the arguments put forth regarding an issue or candidate. In contrast, when people are low in motivation to process information about political issues or candidates (e.g., involvement and personal relevance are low), or when they lack the ability to process systematically (e.g., they are stressed or under time pressure), they may tend to rely on heuristics such as party labels, expert or celebrity endorsements, and source cues such as attractiveness or group membership. For example, a low-motivation or low-capacity voter might oppose a state ballot initiative because Oprah opposes it, support a senator because the letter ("D" or "R") next to the name matches the voter's typical political preferences, or vote for a presidential candidate because their facial features convey an air of competence (see Hall et al., 2009; Todorov et al., 2005).

Political psychologists have identified five broad categories of heuristics that can influence voting behavior: party affiliation, ideological affiliation, endorsements, polls (i.e., consensus information), and candidate appearance (Lau and Redlawsk, 2001). Although in an ideal world, citizens participating in a democratic process would usually think carefully and critically about political information before arriving at conclusion, heuristic processing is thought to guide a substantial portion of political decision-making. For instance, echoing Converse's (1964) observation that the majority of Americans display relatively low levels of political sophistication and knowledge, Mondak (1993) suggested that most voters face a range of pressing everyday concerns that tend to take precedence over political matters, increasing the likelihood that voters will rely on heuristics when processing political information (see also Ledgerwood and Chaiken, 2007). Consistent with this notion, Lau and Redlawsk (2001) found a high rate of heuristic use among individuals participating in a mock presidential election. Using a process-tracing methodology, these researchers were able to track the extent

to which participants accessed different kinds of information about the candidates on a computer: They provided participants with a [p. 261 ↓] list of available types of information (e.g., "Issue Stance," "Past Experience," "Endorsements"), each of which could be opened with a mouse click to display the relevant information, and then recorded which kinds of information participants chose to access. Information from each of Lau and Redlawsk's five political heuristic categories was accessed by over 90 percent of participants. Interestingly, different participants also appeared to prefer different types of heuristics: those higher in political expertise were more likely to use ideology and endorsement heuristics, whereas those lower in expertise were more likely to use candidate appearance heuristics. As the heuristic-systematic model would predict, participants were more likely to use heuristics when their ability to engage in more effortful processing was limited (i.e., when the information environment was made more complex by having the information labels actively scroll past participants on the computer screen rather than remain static).

Lau and Redlawsk's (2001) study suggests that all five categories of heuristics are likely to play a role in a given election; however, some types have been studied more frequently than others. For example, given the prolific use of endorsements for a wide variety of political attitude objects (including everything from local ballot initiatives to presidential candidates), and from a wide variety of endorsers (ranging from political organizations to celebrities), political scientists have been particularly interested in how endorsement heuristics influence political opinions and voting behavior. Using data from a California poll regarding an upcoming election for members of the State Supreme Court, Mondak (1993) showed that endorsements increased voters' willingness to express an opinion and influenced the direction of that opinion when they had relatively little information about the issue. Specifically, respondents were more likely to say a Supreme Court justice should be retained or recalled (rather than choosing "not sure") when told which governor had appointed the justice, and they used their evaluation of the governor to guide their evaluation of the justice in question. In other words, they used the governor's endorsement as a heuristic in forming an attitude toward the associated Supreme Court justice. Consistent with the heuristic-systematic model, this was more likely to occur when respondents had been previously exposed to relatively little media information regarding the justice (thereby limiting their ability to engage in systematic processing) and for respondents scoring higher on a need for cognitive

efficiency measure (designed to tap both motivation and ability to carefully process information).

In Mondak's (1993) study, the heuristic implication of an endorsement from a politician depended on a voter's attitude toward that politician. However, the impact of an endorsement could also depend on the perceived *reliability* of the heuristic for a particular judgmental task; that is, the extent to which a perceiver deems a heuristic to be a valid guide for judgment in a given situation (see Chen and Chaiken, 1999). For example, when considering an environmental issue, a voter might feel that an endorsement from Greenpeace affords a sizeable increase in judgmental confidence, whereas an endorsement from the National Basketball Association does not, despite equivalent evaluations of the two organizations. Indeed, Forehand et al. (2004) found that participants expressed more favorable attitudes toward a hypothetical initiative when it was endorsed by a well-known and issue-relevant source rather than a fictional or issue-irrelevant source. Supporting the heuristic-systematic model's sufficiency principle, this difference emerged in a low motivation context (in which participants expected to justify their preferences about an unimportant and unrelated issue, ballot formatting) but not a high motivation context (in which participants expected to be held accountable for their position on the initiative itself).

Group endorsements can also act to bias systematic information processing about an issue or a candidate. Individuals may be motivated by defense or impression concerns [p. 262 ↓] to agree with an ingroup and disagree with an outgroup, and may therefore process information selectively to arrive at these preferred judgments (Fleming and Petty, 2000). For example, Cohen (2003: Study 4) asked liberal undergraduate students to evaluate a (stereotypically liberal) proposal for a generous federally funded job-training program. Half the participants learned that Democrats (their political ingroup) opposed and Republicans (their political outgroup) supported the program, while half received no information about group endorsement. On average, participants in the no-endorsement condition supported the program, in keeping with their ideological beliefs. However, when participants were told that their ingroup opposed the program, they showed biased processing of the information presented in the proposal, selectively interpreting ambiguous information and selectively attending to unambiguous information to support the ingroup position. As a result, participants in the ingroup-opposed condition were more likely to oppose the program themselves, compared

to participants in the no-information condition. Moreover, the Democratic participants believed that group endorsement influenced the attitudes of other Democrats and (even more strongly) Republicans, but perceived themselves to be relatively unaffected by this information. Thus, consistent with the notion that heuristic processing need not involve intentionality and self-awareness (see Chaiken et al., 1989; Chen and Chaiken, 1999), it seems likely that people are unaware of the extent to which group endorsements bias their thinking about an issue. This may tend to exacerbate political conflicts: whereas Democrats and Republicans might agree on a policy in the absence of endorsement information, merely attaching a party label to a proposal can distort information processing and lead partisans to adopt divergent positions. Interestingly, then, bipartisan proposals may be particularly likely to gain public support not only because their actual content may better address the political goals of both groups, but also because the absence of a link to a particular party may help to promote more open-minded information processing.

Conclusion

Looking back, we see the heuristic-systematic model as very much a product of its historical context, building on theories both within the attitudes domain and outside of it, and developing beyond the study of basic social psychological processes to shed light on important and relevant social issues. To us, this illustrates the benefit of working in an area with such a long and cumulative history that both influences and draws from other psychological and related social-science disciplines. In coming years, we hope that the field continues to develop the heuristic-systematic model in concert with other dual-process theories, drawing from the research that has already been done to influence that which is yet to come.

References

- Allport, G. (1954) *The Nature of Prejudice* . Reading, MA: Addison-Wesley.

Anderson, R.C. and Pichert, J.W. (1978) Recall of previously unrecallable information following a shift in perspective . *Journal of Verbal Learning and Verbal Behavior* , 17, 1–12. <http://dx.doi.org/10.1016/S0022-5371%2878%2990485-1>

Axsom, D., Yates, S.M. and Chaiken, S. (1987) Audience response as a heuristic cue in persuasion . *Journal of Personality and Social Psychology* , 53, 30–40. <http://dx.doi.org/10.1037/0022-3514.53.1.30>

Bargh, J.A. (1994) The Four Horsemen of automaticity: Awareness, efficiency, intention, and control in social cognition . In R.S. Wyer, Jr., ed. and T.K. Srull (eds), *Handbook of Social Cognition* , 2nd Edition, pp. 1–40. Hillsdale, NJ: Erlbaum.

Baumeister, R.F. and Leary, M.R. (1995) The need to belong: Desire for interpersonal attachments as a fundamental human motivation . *Psychological Bulletin* , 117, 497–529. <http://dx.doi.org/10.1037/0033-2909.117.3.497>

Bem, D.J. (1972) Self-perception theory . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 6, 1–62. New York: Academic Press.

Brewer, M.B. (1988) A dual-process model of impression formation . In T. Srull, ed. and R. Wyer (eds), *Advances in Social Cognition* , 1, 1–36. Hillsdale, NJ: Erlbaum.

Brewer, N. and Hupfeld, R.M. (2004) Effects of testimonial inconsistencies and witness group identity on mock-juror judgments . *Journal of Applied Social Psychology* , 34, 493–513. <http://dx.doi.org/10.1111/j.1559-1816.2004.tb02558.x>

Brewer, W.F. and Treyens, J.C. (1981) Role of schemata in memory for places . *Cognitive Psychology* , 13, 207–230. <http://dx.doi.org/10.1016/0010-0285%2881%2990008-6>

Bruner, J.S. (1957) On perceptual readiness . *Psychological Review* , 64, 123–152. <http://dx.doi.org/10.1037/h0043805>

Chaiken, S. (1980) Heuristic versus systematic information processing and the use of source versus message cues in persuasion . *Journal of Personality and Social Psychology* , 39, 752–766. <http://dx.doi.org/10.1037/0022-3514.39.5.752>

Chaiken, S. (1987) The heuristic model of persuasion . In M.P. Zanna, ed. , J.M. Olson, ed. and C.P. Herman (eds), Social Influence: The Ontario Symposium , 5, 3–39. Hillsdale, NJ: Erlbaum.

Chaiken, S., Axsom, D., Liberman, A. and Wilson, D. (1992) Heuristic processing of persuasive messages: Chronic and temporary sources of rule accessibility . Unpublished manuscript, New York University.

Chaiken, S. and Eagly, A.H. (1976) Communication modality as a determinant of message persuasiveness and message comprehension . Journal of Personality and Social Psychology , 34, 605–614. <http://dx.doi.org/10.1037/0022-3514.34.4.605>

Chaiken, S. and Eagly, A.H. (1983) Communication modality as a determinant of persuasion: The role of communicator salience . Journal of Personality and Social Psychology , 45, 241–256. <http://dx.doi.org/10.1037/0022-3514.45.2.241>

Chaiken, S., Giner-Sorolla, R. and Chen, S. (1996) Beyond accuracy: multiple motives in heuristic and systematic processing . In P.M. Gollwitzer, ed. and J.A. Bargh (eds), The Psychology of Action: Linking Motivation and Cognition to Action , pp. 553–578. New York: Guilford Press.

Chaiken, S., Liberman, A. and Eagly, A.H. (1989) Heuristic and systematic information processing within and beyond the persuasion context . In J.S. Uleman, ed. and J.A. Bargh (eds), Unintended Thought , pp. 212–252. New York: Guilford Press.

Chaiken, S. and Maheswaran, D. (1994) Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment . Journal of Personality and Social Psychology , 66, 460–473. <http://dx.doi.org/10.1037/0022-3514.66.3.460>

Chaiken, S., ed. and Trope, Y. (eds) (1999) Dual-Process Theories in Social Psychology . New York: Guilford Press.

Chen, S. and Chaiken, S. (1999) The heuristic-systematic model in its broader context . In S. Chaiken, ed. and Y. Trope (eds), Dual-process Theories in Social Psychology , pp. 73–96. New York: Guilford Press.

Chen, S., Shechter, D. and Chaiken, S. (1996) Getting at the truth or getting along: Accuracy- versus impression-motivated heuristic and systematic processing . *Journal of Personality and Social Psychology* , 71, 262–275. <http://dx.doi.org/10.1037/0022-3514.71.2.262>

Cohen, G.L. (2003) Party over policy: The dominating impact of group influence on political beliefs . *Journal of Personality and Social Psychology* , 85, 808–822. <http://dx.doi.org/10.1037/0022-3514.85.5.808>

Converse, P.E. (1964) The nature of belief systems in mass publics . In D.E. Apter (ed.), *Ideology and Discontent* , pp. 206–261. New York: Free Press.

Darke, P.R., Chaiken, S., Bohner, G., Einwiller, S., Erb, H.-P. and Hazlewood, J.D. (1998) Accuracy motivation, consensus information, and the law of large numbers: Effects on attitude judgment in the absence of argumentation . *Personality and Social Psychology Bulletin* , 24, 1205–1215. <http://dx.doi.org/10.1177/01461672982411007>

De Dreu, C.K.W. (2003) Time pressure and closing of the mind in negotiation . *Organizational Behavior and Human Decision Processes* , 91, 280–295. <http://dx.doi.org/10.1016/S0749-5978%2803%2900022-0>

De Dreu, C.K.W. (2004) Motivation in negotiation: A social psychological analysis . In M. Gelfand, ed. and J.M. Brett (eds), *Handbook of Negotiation and Culture* , pp. 114–135. Palo Alto, CA: Stanford University Press.

De Dreu, C.K.W., Koole, S. and Oldersma, F.L (1999) On the seizing and freezing of negotiator inferences: Need for cognitive closure moderates the use of heuristics in negotiation . *Personality and Social Psychology Bulletin* , 25, 348–362. <http://dx.doi.org/10.1177/0146167299025003007>

De Dreu, C.K.W., Koole, S.L. and Steinle, S. (2000) Unfixing the fixed pie: A motivated information-processing approach to integrative negotiation . *Journal of Personality and Social Psychology* , 79, 975–987. <http://dx.doi.org/10.1037/0022-3514.79.6.975>

Deutsch, M. and Gerard, H.B. (1955) A study of normative and informational social influences upon individual judgment . *Journal of Abnormal Social Psychology* , 51, 629–636. <http://dx.doi.org/10.1037/h0046408>

Devine, P.G. (1989) Stereotypes and prejudice: Their automatic and controlled components . *Journal of Personality and Social Psychology* , 56, 5–18. <http://dx.doi.org/10.1037/0022-3514.56.1.5>

Ditto, P.H. and Lopez, D.F. (1992) Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions . *Journal of Personality and Social Psychology* , 63, 568–584. <http://dx.doi.org/10.1037/0022-3514.63.4.568>

Eagly, A.H. and Chaiken, S. (1993) *The Psychology of Attitudes* . Belmont, CA: Wadsworth.

Erb, H.P., Bohner, G., Schmalzle, K. and Rank, S. (1998) Beyond conflict and discrepancy: cognitive bias in minority and majority influence . *Personality and Social Psychology Bulletin* , 24, 620–633. <http://dx.doi.org/10.1177/0146167298246006>

Fazio, R.H. and Towles-Schwen, T. (1999) The MODE model of attitude-behavior processes . In S. Chaiken, ed. and Y. Trope (eds), *Dual-process Theories in Social Psychology* , pp. 97–116. New York: Guilford.

Fiske, S.T. (2002) Five core social motives, plus or minus five . In S.J. Spencer, ed. , S. Fein, ed. , M.P. Zanna, ed. and J. Olson (eds), *Motivated Social Perception: The Ontario Symposium* , 9, 233–246. Mahwah, NJ: Erlbaum.

Fiske, S.T., Lin, M.H. and Neuberg, S.L. (1999) The Continuum Model: Ten years later . In S. Chaiken, ed. and Y. Trope (eds), *Dual-process Theories in Social Psychology* , pp. 231–254. New York: Guilford Press.

Fiske, S.T. and Linville, P.W. (1980) What does the schema concept buy us? *Personality and Social Psychology Bulletin* , 6, 543–557. <http://dx.doi.org/10.1177/014616728064006>

Fiske, S.T. and Taylor, S.E. (2008) Social Cognition: From Brains to Culture . New York: McGraw-Hill.

Fleming, M.A. and Petty, R.E. (2000) Identity and persuasion: An elaboration likelihood approach . In D.J. Terry, ed. and M.A. Hogg (eds), Attitudes, Behavior, and Social Context: The Role of Norms and Group Membership , pp. 171–199. Mahwah, NJ: Lawrence Erlbaum Associates.

Forehand, M., Gastil, J. and Smith, M.A. (2004) Endorsements as voting cues: Heuristic and systematic processing in initiative elections . Journal of Applied Social Psychology , 34, 2215–2233. <http://dx.doi.org/10.1111/j.1559-1816.2004.tb01974.x>

ForsterLee, R., ForsterLee, L., Horowitz, I.A. and King, E. (2006) The effects of defendant race, victim race, and juror gender on evidence processing in a murder trial . Behavioral Sciences and the Law , 24, 179–198. <http://dx.doi.org/10.1002/bls.675>

Gawronski, B. and Bodenhausen, G.V. (2006) Associative and prepositional processes in evaluation: An integrative review of implicit and explicit attitude change . Psychological Bulletin , 132, 692–731. <http://dx.doi.org/10.1037/0033-2909.132.5.692>

Gilbert, D.T. (1989) Thinking lightly about others: Automatic components of the social inference process . In J.S. Uleman, ed. and J.A. Bargh (eds), Unintended Thought , pp. 189–211. New York: Guilford.

Giner-Sorolla, R. and Chaiken, S. (1997) Selective use of heuristic and systematic processing under defense motivation . Personality and Social Psychology Bulletin , 23, 84–97. <http://dx.doi.org/10.1177/0146167297231009>

Giner-Sorolla, R., Chaiken, S. and Lutz, S. (2002) Validity beliefs and ideology can influence legal case judgments differently . Law and Human Behavior , 26, 507–526. <http://dx.doi.org/10.1023/A:1020251921659>

Hall, C.C, Goren, A., Chaiken, S. and Todorov, A. (2009) Shallow cues with deep effects: Trait judgments from faces and voting decisions . In E. Borgida, ed. , J.L. Sullivan, ed. , and C.M. Federico (eds), The Political Psychology of Democratic

Citizenship , pp. 73–99. Oxford: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780195335453.003.0004>

Harris, P.R. and Napper, L. (2005) Self-affirmation and the biased processing of health-risk information . *Personality and Social Psychology Bulletin* , 31, 1250–1263. <http://dx.doi.org/10.1177/0146167205274694>

Heider, F. (1958) *The Psychology of Interpersonal Relations* . Hoboken, NJ: Wiley. <http://dx.doi.org/10.1037/10628-000>

Higgins, E.T. (1989) Knowledge accessibility and activation: Subjectivity and suffering from unconscious sources . In J.S. Uleman, ed. and J.A. Bargh (eds), *Unintended Thought* , pp. 75–115. New York: Guilford Press.

Higgins, E.T., King, G.A. and Mavin, G.H. (1982) Individual construct accessibility and subjective impressions and recall . *Journal of Personality & Social Psychology* , 43, 35–47. <http://dx.doi.org/10.1037/0022-3514.43.1.35>

Higgins, E.T. and McCann, C.D. (1984) Social encoding and subsequent attitudes, impressions, and memory: ‘Context-driven’ and motivational aspects of processing . *Journal of Personality and Social Psychology* , 47, 26–39. <http://dx.doi.org/10.1037/0022-3514.47.1.26>

Howard, M.V.A, Brewer, N. and Williams, K.D. (2006) How processing resources shape the influence of stealing thunder on mock-juror verdicts . *Psychiatry, Psychology and Law* , 13, 60–66. <http://dx.doi.org/10.1375/pplt.13.1.60>

Jepson, C. and Chaiken, S. (1990) Chronic issue-specific fear inhibits systematic processing of persuasive communications . *Journal of Social Behavior and Personality* , 5, 61–84.

Johnson, B.T. and Eagly, A.H. (1989) The effects of involvement on persuasion: A meta-analysis . *Psychological Bulletin* , 106, 290–314. <http://dx.doi.org/10.1037/0033-2909.106.2.290>

Jones, E.E. and Davis, K.E. (1965) From acts to dispositions: The attribution process in social psychology . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 2, 219–266. New York: Academic Press.

Katz, D. (1960) The functional approach to the study of attitudes . *Public Opinion Quarterly* , 24, 163–204. <http://dx.doi.org/10.1086/266945>

Kelley, H.H. (1972) Causal Schemata and the Attribution Process . New York: General Learning Press.

Kelley, H.H. (1973) The processes of causal attribution . *American Psychologist* , 28, 107–128. <http://dx.doi.org/10.1037/h0034225>

Kiesler, C.A. (1971) *The Psychology of Commitment-Experiments Linking Behavior to Belief* . San Diego, CA: Academic Press.

Kohler, W. (1930) *Gestalt Psychology* . London: Bell.

Lau, R.R. and Redlawsk, D.P. (2001) An experimental study of information search, memory, and decision making during a political campaign . In J. Kuklinski (ed.) *Citizens and Politics: Perspectives from Political Psychology* pp. 136–159. New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511896941.008>

Ledgerwood, A. and Chaiken, S. (2007) Priming us and them: Automatic assimilation and contrast in group attitudes . *Journal of Personality and Social Psychology* , 93, 940–956. <http://dx.doi.org/10.1037/0022-3514.93.6.940>

Ledgerwood, A., Chaiken, S., Gruenfeld, D.H. and Judd, C.M. (2006) Changing minds: Persuasion in negotiation and conflict resolution . In M. Deutsch, ed. , P.T. Coleman, ed. , and E.C. Marcus (eds), *The Handbook of Conflict Resolution: Theory and Practice* , 2nd Edition, pp. 455–485. San Francisco: Jossey-Bass.

Ledgerwood, A., Mandisodza, A.N., Jost, J.T. and Pohl, M.J. (in press) Working for the system: Motivated defense of meritocratic beliefs . *Social Cognition* .

Lerner, J.S. and Tetlock, P.E. (1999) Accounting for the effects of accountability . Psychological Bulletin , 125, 255–275. <http://dx.doi.org/10.1037/0033-2909.125.2.255>

Liberman, A. and Chaiken, S. (1992) Defensive processing of personally relevant health messages . Personality and Social Psychology Bulletin , 18, 669–679. <http://dx.doi.org/10.1177/0146167292186002>

Lord, C.G., Ross, L. and Lepper, M.R. (1979) Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence . Journal of Personality and Social Psychology , 37, 2098–2109. <http://dx.doi.org/10.1037/0022-3514.37.11.2098>

Mackie, D.M. (1987) Systematic and nonsystematic processing of majority and minority persuasive communications . Journal of Personality and Social Psychology , 53, 41–52. <http://dx.doi.org/10.1037/0022-3514.53.1.41>

Maheswaran, D. and Chaiken, S. (1991) Promoting systematic processing in low motivation settings: Effect of incongruent information on processing and judgment . Journal of Personality and Social Psychology , 61, 13–25. <http://dx.doi.org/10.1037/0022-3514.61.1.13>

Maheswaran, D., Mackie, D.M. and Chaiken, S. (1992) Brand name as a heuristic cue: The effects of task importance and expectancy confirmation on consumer judgments . Journal of Consumer Psychology , 1, 317–336. <http://dx.doi.org/10.1016/S1057-7408%2808%2980058-7>

Marcus, G.E., Neumann, W.R. and MacKuen, M.B. (2000) Affective Intelligence and Political Judgment . Chicago: University of Chicago Press.

Martin, R., Hewstone, M. and Martin, P.Y. (2007) Systematic and heuristic processing of majority and minority endorsed messages: The effects of varying ‘levels of orientation’ and outcome relevance on attitude and message processing . Personality and Social Psychology Bulletin , 33, 43–56. <http://dx.doi.org/10.1177/0146167206294251>

McGuire, W.J. (1969) The nature of attitudes and attitude change . In G. Lindzey, ed. and E. Aronson (eds), *Handbook of Social Psychology* , 2nd Edition, 3, 136–314. Reading, MA: Addison-Wesley.

Mondak, J.J. (1993) Public opinion and heuristic processing of source cues . *Political Behavior* , 15, 167–192. <http://dx.doi.org/10.1007/BF00993852>

Moskowitz, G.B., Skurnik, L and Galinsky, A.D. (1999) The history of dual-process notions, and the future of pre-conscious control . In S. Chaiken, ed. and Y. Trope (eds), *Dual-process Theories in Social Psychology* , pp. 12–36. New York: Guilford Press.

Newman, B.I. and Perloff, R.M. (2004) Political marketing: theory, research, and applications . In L.L. Kaid (ed.), *Handbook of Political Communication Research* , pp. 17–43. Mahwah, NJ: Erlbaum.

Pallak, M.S., Mueller, M., Dollar, K. and Pallak, J. (1972) Effect of commitment on responsiveness to an extreme consonant communication . *Journal of Personality and Social Psychology* , 23, 429–436. <http://dx.doi.org/10.1037/h0033363>

Petty, R.E. and Cacioppo, J.T. (1984) Source factors and the elaboration likelihood model of persuasion . *Advances in Consumer Research* , 11, 668–672.

Petty, R.E., Cacioppo, J.T. and Goldman, R. (1981) Personal involvement as a determinant of argument-based persuasion . *Journal of Personality and Social Psychology* , 41, 847–855. <http://dx.doi.org/10.1037/0022-3514.41.5.847>

Petty, R.E. and Wegener, D.T. (1999) The elaboration likelihood model: Current status and controversies . In S. Chaiken, ed. and Y. Trope (eds), *Dual-process Theories in Social Psychology* , pp. 37–72. New York: Guilford.

Petty, R.E., Wells, G.L. and Brock, T.C. (1976) Distraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification . *Journal of Personality and Social Psychology* , 34, 874–884. <http://dx.doi.org/10.1037/0022-3514.34.5.874>

Ratneshwar, S. and Chaiken, S. (1991) Comprehension's role in persuasion: The case of its moderating effect on the persuasive impact of source cues . *Journal of Consumer Research* , 18, 52–62. <http://dx.doi.org/10.1086/209240>

Scott, C.G. (1996) Understanding attitude change in developing effective substance abuse prevention programs for adolescents . *The School Counselor* , 43, 187–195.

Sherif, M. and Cantril, H. (1947) *The Psychology of Ego-involvements: Social Attitudes and Identifications* . New York: Wiley. <http://dx.doi.org/10.1037/10840-000>

Sherif, M. and Hovland, C.I. (1961) *Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change* . New Haven, CT: Yale University Press.

Simon, H.A. (1976). *Administrative Behavior* , 3rd Edition. New York: Free Press.

Smith, E.R. and DeCoster, J. (2000) Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems . *Personality and Social Psychology Review* , 4, 108–131. http://dx.doi.org/10.1207/S15327957PSPR0402_01

Smith, M.B., Bruner, J.S. and White, R.W. (1956) *Opinions and Personality* . New York: Wiley.

Stroebe, W. and Diehl, M. (1988) When social support fails: Supporter characteristics in compliance-induced attitude change . *Personality and Social Psychology Bulletin* , 14, 136–144. <http://dx.doi.org/10.1177/0146167288141014>

Taylor, S.E. (1975) On inferring one's attitudes from one's behavior: Some delimiting conditions . *Journal of Personality and Social Psychology* , 31, 126–131. <http://dx.doi.org/10.1037/h0076246>

Taylor, S.E. and Crocker, J. (1981) Schematic bases of social information processing . In E.T. Higgins, ed. , C.P. Herman, ed. , and M.P. Zanna (eds), *Social Cognition: The Ontario Symposium* , 1, pp. 89–134. Hillsdale, NJ: Erlbaum.

Tesser, A. and Campbell, J. (1983) Self-definition and self-evaluation maintenance . In J. Suis, ed. and A. Greenwald (eds), Social Psychological Perspectives on the Self , 2, 1–31. Hillsdale, NJ: Erlbaum.

Tetlock, P.E. (1989) Structure and function in political belief systems . In A.R. Pratkanis, ed. , S.J. Breckler, ed. , and A.G. Greenwald (eds), Attitude Structure and Function , pp. 129–151. Hillsdale, NJ: Erlbaum.

Tetlock, P.E., Skitka, L. and Boettger, R. (1989) Social and cognitive strategies for coping with accountability: Conformity, complexity, and bolstering . *Journal of Personality and Social Psychology* , 57, 632–640. <http://dx.doi.org/10.1037/0022-3514.57.4.632>

Thompson, L.L. and Hastie, R. (1990) Social perception in negotiation . *Organizational Behavior and Human Decision Processes* , 47, 98–123. <http://dx.doi.org/10.1016/0749-5978%2890%2990048-E>

Todorov, A., Mandisodza, A.N., Goren, A. and Hall, C.C. (2005) Inferences of competence from faces predict election outcomes . *Science* , 308, 1623–1626. <http://dx.doi.org/10.1126/science.1110589>

Tversky, A. and Kahneman, D. (1974) Judgment under uncertainty: Heuristics and biases . *Science* , 185, 1124–1131. <http://dx.doi.org/10.1126/science.185.4157.1124>

Webster, D.M. and Kruglanski, A.W. (1994) Individual differences in need for cognitive closure . *Journal of Personality and Social Psychology* , 67, 1049–1062. <http://dx.doi.org/10.1037/0022-3514.67.6.1049>

Werner, C.M., Stroll, R., Birch, P. and White, P.H. (2002) Clinical validation and cognitive elaboration: Signs that encourage sustained recycling . *Basic and Applied Social Psychology* , 24, 185–203.

Wood, W., Kallgren, C.A. and Preisler, R.M (1985) Access to attitude-relevant information in memory as a determinant of persuasion: The role of message attributes . *Journal of Experimental Social Psychology* , 21, 73–85. <http://dx.doi.org/10.1016/0022-1031%2885%2990007-1>

Yates, J. (1985) The content of awareness is a model of the world . Psychological Review , 92, 249–284. <http://dx.doi.org/10.1037/0033-295X.92.2.249>

Ziegler, R., von Schwichow, A. and Diehl, M. (2005) Matching the message source to attitude functions: Implications for biased processing . Journal of Experimental Social Psychology , 41, 645–653. <http://dx.doi.org/10.1016/j.jesp.2004.12.002>

Zuckerman, A. and Chaiken, S. (1998) A heuristic-systematic processing analysis of the effectiveness of product warning labels . Psychology and Marketing , 15, 621–664. <http://dx.doi.org/10.1002/%28SICI%291520-6793%28199810%2915:7%3C621::AID-MAR2%3E3.0.CO;2-H>

<http://dx.doi.org/10.4135/9781446249215.n13>

Handbook of Theories of Social Psychology: Volume 1

The Continuum Model and the Stereotype Content Model

Contributors: Susan T. Fiske

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "The Continuum Model and the Stereotype Content Model"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n14>

Print pages: 267-289

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n14>

[p. 267 ↓]

Chapter 13: The Continuum Model and the Stereotype Content Model

Susan T. Fiske, ed.

Abstract

Two models address respectively processes and contents of social categories in impression formation. In the continuum model (CM), people default to category-based processes to form impressions of others, so they often respond rapidly, captured by schema-triggered affect, stereotypic associations, and discriminatory responses.

According to the CM, adequate information and appropriate motivation can override categories and encourage more individuating processes. What are the default categories that capture social cognition? In the stereotype content model (SCM), two fundamental dimensions describe clusters of apparently universal categories for interpersonal and intergroup perception. The warmth dimension divides people into friends and foes, while the competence dimension divides people by status. Warmth-by-competence combinations identify (a) the ingroup plus its allies, seen as both warm and competent, eliciting pride (e.g., middle-class people), contrasting with three outgroup types; (b) extreme, disgust-inducing groups, allegedly neither warm nor competent (e.g., homeless people); (c) ambivalent, pitied groups, allegedly warm but incompetent (e.g., older people); and (d) ambivalent, envied groups, allegedly competent but not warm (e.g., rich people). Distinct active and passive discrimination target each quadrant, along with preliminary evidence of neural correlates. Not all categories are

equivalent; CM and SCM principles describe both usage and patterns of content with policy implications.

Categorically Captured: The Continuum Model and the Stereotype Content Model

Puzzle #1: At my high-school best-friend's party, over a pounding Motown beat, one random dance partner asked me, "What sign are you?" Prior to that, he knew me only as a white girl about his age, probably attending the local private school. I knew him only as a black guy about my age, presumably from the neighborhood, though not from my school. When I answered, between beats, "Leo," he turned and walked away, in the middle of the song. I was less hurt than merely puzzled as to what that category meant for him.

[p. 268 ↓] People must perceive with the aid of categories; as Allport noted, "[O]rderly living depends upon it" (1954: 19). Consequently, our social categories capture our impressions. We cannot help noticing a person's gender, race, and age within a fraction of a second (see Fiske and Taylor, 2008, for a review). We also instantly perceive cues to social class, status, and some predispositions such as extraversion (though astrological signs do not appear among the most immediately visible cues). These accessible expectations shape our responses from the first moments. Two of my theories, developed with a little help from my friends, describe, respectively, the processes and the contents of our social categories. And they solve some puzzles along the way, including how people use categories such as "Leo."

A Personal Tale of Two Theories

I admit to being obsessed by categories. Attending Hyde Park's University of Chicago Laboratory Schools – later to enroll Sasha and Malia Obama – meant even then encountering an abundantly multiracial and economic diversity of students, from nursery school onward. We played at each other's houses and knew each other's families. We collaborated and competed in class, clubs, and sports. We danced and dated. Not that

we failed to notice race, class, or gender, but we had additional impressions to guide our play, our projects, and our schoolwork. Nevertheless came Puzzle #2: In the clique-infested halls of middle school, former elementary-school friends suddenly separated into identity groups based on income, race, and ethnicity, as well as intellectual and athletic aspirations. The cafeteria seating clusters would be familiar to anyone attending middle or high school. However, being a small community, we had known each other too long to view each other as aliens. How did we live in these two worlds of identity politics and deep familiarity?

Graduation later scattered us, and my own college experience of Boston shocked me into realizing that Chicago's Hyde Park possessed a distinct microclimate. The rest of the world seemed to operate along stricter racial and economic divides than I knew. Puzzle #3: What made the difference?

Transition to graduate school in the 1970s, when issues of race, class, gender, and group conflict dominated late-night discussions in dorms, apartments, and – yes – my communal house. One debated question was why a local doorman kept confusing me with Kathleen Kennedy. Apart from gender, age, race, and class, why would he mix us up? (Puzzle #4). With perfect timing, my graduate advisor, Shelley Taylor, invited me on board some experiments studying social categorization; I leapt at the opportunity. As some of our first studies indicated (Taylor et al., 1978), people code each other by gender and race, so much so that they confuse other people within category (say, blacks with other blacks, and whites with other whites) more than between categories (rarely confusing blacks with whites or vice versa).

In one of our studies, Harvard undergraduates watched a prerecorded brainstorming conversation, in which each person offered suggestions about how to publicize a play. Afterward, in a surprise recognition task, participants were more likely to make with-in-category errors than between-category ones, as predicted. That is, they misremembered some of a black person's suggestions as coming from a different black person. Ditto for whites with other whites, as well as men with other men and women with other women. This was some of the first evidence for the role of basic social categories in social cognition (later much replicated). Puzzle #4 again: Why wouldn't people pay more attention, beyond race and gender?

And now for something completely different: in graduate school, as is typical, I worked as a teaching assistant for several semesters. I began to observe a phenomenon regarding my impressions of my students, namely that even when I could not remember someone's [p. 269 ↓] full identity, I always knew how I felt about the person. How did these impressions form separately, the slippery facts and the reliable likability? This Puzzle #5 fit another research project, completely different from the category-oriented one, so it requires a little background.

At this time, I was also working with Reid Hastie on impression formation according to Norman Anderson's (1981) information integration model of evaluative impressions (elaborated in the next section). His theory predicted how perceivers combine individual cues to form an overall evaluation. Our study examined whether such evaluations and the relevant memory follow the same or different cognitive processes. For Elizabeth Drebens senior thesis (Drebens et al., 1979), Harvard undergraduates read a series of positive, negative, and neutral sentences describing a person's behavior (e.g., "Alan bought groceries for an elderly lady next door who was ill"; "Alan pressed the button and waited for the elevator to come"). People apparently formed and updated their impressions online, as they received each cue, but their final impressions did not depend on memory for the behaviors. This study provided some of the first evidence for the independence of impressions and memory in the encoding of individual cues (Hastie and Park, 1986). At each moment, as they receive each piece of information, people immediately know how they feel about other people, continually updating that impression, but later they cannot necessarily produce the relevant evidence. Ever known exactly how you feel about somebody, without recalling any of the reason for those feelings? Puzzle #5 again. Now turn to a more formal statement of our models that ultimately help explain these puzzles.

Intellectual History of the Models

Why were the Taylor and Hastie studies precedent-setting? Why was it so exciting to be present at their inception? First, some intellectual context explains why, and then the next sections describe two Fiske et al. models that emerged to address some of social cognition's continuing puzzles.

Intellectual Context: Birth of Social Cognition

Both these graduate-school collaborations joined the opening salvos of the social cognitive revolution in psychological sciences. The first half of the twentieth century had seen psychology dominated by learning theories that valorized stimulus and response; that is, directly observable conditions and behaviors (Boring, 1950). The mind – intervening mental processes of organisms (from lab rats to undergraduates) – had been off-limits, judged scientifically irrelevant, unimportant, or even suspect. For example, Guthrie famously made fun of Tolman's maze-rat "buried in thought" and never choosing (1935: 172).

At mid-century, this tyranny was overthrown by the cognitive revolution, advocating the study of humans as information processors, with minds oriented toward encoding, representation, retrieval, and response. The computer metaphor informed studies of attention, memory, and problem solving. Cognitive psychology flourished. Before and during the cognitive revolution, meanwhile, offices on psychology departments' other floors had long opposed the field's prior domination by learning-theory diktats: social psychologists had always maintained their focus on indirectly observable mental constructs, such as attitudes and stereotypes (for an overview, see Fiske and Taylor, 2008). With the cognitive revolution, some social psychologists leapt into familiar territory (impressions) seen from a new angle. The new ideas turned out to have exciting implications for how people make sense, not just of their inanimate world, but also of each other.

Into this heady intellectual era, my research launched two separate trajectories. [p. 270 ↓] My graduate work with Taylor on social categories continued in studies of category salience (Taylor and Fiske, 1978). Among them were studies of how solo (or token) members of under-represented groups suffer exaggerated, category-constrained judgments. The explanatory potential of categories appeared crucial.

At the same time, my dissertation picked up the Anderson information integration theory. With its focus on the isolated impact of each separate social cue, I hoped

eventually to design studies that would quantify the impact of race, gender, and other impermissible categories on impression formation, relative to the other information available. My dissertation applied the information integration formulas to individual behaviors because they seemed to constitute a more elementary unit than personality traits, let alone social categories. In my rash cross-country phone consult, Norman Anderson firmly advised me not to use the rich but messy behavioral stimuli proposed.

Undaunted, my dissertation project did examine behaviors, namely the disproportionate weight of negative and extreme behaviors, relative to positive and moderate ones (Fiske, 1980). In a painstaking 16×16 Latin square, participants viewed images of 16 separate young men, each depicted at one of four levels of sociability (from isolation to full involvement at a picnic) and at four levels of civic responsibility (from rejection to promotion of a petition against child pornography). Participants' looking-time recorded for each slide, they rated each person's likability in turn. As predicted, the theoretically derived impression weights for each behavior closely mirrored people's online visual attention to each behavior. The study linked these computed impression weights to observable behavior, showing that both negative and extreme behaviors independently elicit both weight and attention. The study almost won the Society of Experimental Social Psychology (SESP) dissertation award, and helped me land a job at that information-processing Mecca, Carnegie-Mellon University.

Just as my two separate lines of research seemed to be launched, nearby Ohio State University invited me for a colloquium in my first year as an assistant professor. An astute but now-nameless graduate student listened to the account of my dissertation research, knowing about my other work on social categories, and asked the life-changing question: "how can you do both?" The apparent contradiction struck me all of a heap: on one hand, the elemental, piecemeal, information-integration approach famously denied any interaction among the elements of an impression. Mathematically, it proposes a weighted averaging model, wherein each cue possesses: s_i

, a set value on a likability scale (e.g., on a 100-point scale, how unlikable it is to reject friends at a picnic), and w_i

, a weight reflecting its relative importance in the overall impression (e.g., social rejection might matter more than civic responsibility to one's overall likability). Regardless of its other algebraic assumptions, the model stated clearly that the elements would not affect each other's evaluations (e.g., the sociability of an irresponsible person is just as likable as the sociability of a responsible person). From a mathematical perspective, this meant that the individual elements could not interact or multiply. Hence, the terms *elemental* or *piecemeal* both describe this theory of impression formation.

On the other hand, category theories (or schema theories, as they were often known) had an intrinsically interactive flavor. For example, two landmark studies in my professional cohort established the importance of such processes. Nancy Cantor (e.g., Cantor and Mischel, 1977) showed that category prototypes intrude on related trait judgments, so that conceptually related but not presented items bias recognition memory. Hazel Markus (e.g., 1977) showed that the most central parts of people's self-concepts operate like memory schemas organizing their own traits. These and other early social cognition studies built on older Gestalt theories that argued [p. 271 ↓] for configural features of perception. Applied to person perception, the *configural* approach implies that the meaning of any given cue (e.g., intelligent) depends on the context of its surrounding traits. In Solomon Asch's famous (1946) demonstration, a warm intelligent person is wise, whereas a cold intelligent person is sly.

The Gestalt meaning-change versus Andersonian fixed-scale-value perspectives inspired heated debates (Hamilton and Zanna, 1974). Because one of the protagonists taught at Ohio State (Ostrom, 1977), perhaps the student's pivotal question reflected local preoccupations. The question certainly catalyzed mine. That puzzle (#6, but who's counting) began to resolve some of the earlier ones, as a result of motivating the continuum model.

A Category Process Model: The Continuum Model of Impression Formation

The wordy subtitle to our eventual main theoretical statement provides a précis of the continuum model (CM): from category-based to individuating processes, influence of information and motivation on attention and interpretation (Fiske and Neuberg, 1990). The CM resulted directly from the Ohio State graduate student pointing out the apparent contradiction between my two lines of research, examining both elemental and categorical processes. On reflection, choosing one process, to the exclusion of the other, made no sense. Too much research already demonstrated the viability of each type of process using different paradigms. My hunch was that people would do both, but under different circumstances that not coincidentally reflected the respective paradigms used to establish the distinct processes. The theoretical task was to demonstrate how people do which one when.

Our team's initial efforts were more empirical than explicitly theoretical, anticipated in a 1982 grant proposal, whose overall goal was to contrast category-based processing, characteristic of stereotyping, with individuating processes, characteristic of piecemeal impressions. Nevertheless, the studies test the model that eventually evolved out of that work, so I will start here with the subsequently official CM hypotheses (Fiske and Neuberg, 1990):

- Category-based processes take priority over attribute-based process, both because people start with categories and because they stop at categories, if the categories work well enough. Categories are instantly available, convenient, and linked to ready responses – stereotypes, prejudices, and discriminatory tendencies. If the target individual adequately fits the category prototype, then perceivers should not bother collecting and analyzing details about the target.
- Perceivers progress along the continuum, from category-based processes, to recategorization, to more attribute-based processes, depending on adequate information fit. Although captured by categories, perceivers are no fools, and if the category obviously does not fit (Obama does not appear to be a

stereotypic black man), people first recategorize by picking another category (maybe he is a Muslim), subtyping (maybe he is a black professional), making a new category (biracial, cosmoethnic), or relating him to self (he's a citizen of the world, like me). If none of those compromises fit, then people construct a more deliberate, attribute-by-attribute, piecemeal impression (Hawaiian birth, Kansas upbringing, anthropologist mother, Kenyan father, Indonesian elementary school, Columbia, Harvard, Chicago community organizer, Illinois legislator...).

- Attention mediates the use of attribute information, beyond the initial category. Each step requires progressively more attentional resources, and focus of attention determines information use. For example, attentional focus on category-consistent information tends to reconfirm the category, whereas attention to inconsistent information allows category disconfirmation.
- Motivation, especially interdependence, moderates progress along the continuum. People accept initial, default categories when they have no motivation to expend resources examining further. When accuracy especially matters (because they depend on the other person, for example, or because they personally prioritize accurate impressions), people are motivated to try harder.
- Attention also mediates the effects of motivation on impressions. Accuracy-motivated people focus on apparent inconsistencies, in an effort to resolve them by changing their meaning to fit the category or revising their impression. People also can be motivated to confirm their stereotypes, and in that case, they might focus on category-confirming attributes.

Antecedents of each hypothesis came from prior theorists. A version of the first hypothesis had long ago appeared in Asch's (1946) paper that favored Gestalt configurational processes over more elemental (algebraic) processes, although we were among the first applying modern information-processing theories to these ideas. Information fit, as a moderator, built on person-memory research that contrasted congruent and incongruent information (e.g., Hastie et al., 1980). Motivation, as a moderator, built on dissatisfactions with the cold-cognition approaches (Clark and Fiske, 1982). And the mediator in both cases, attention, built on my dissertation and prior graduate work on salience (Taylor and Fiske, 1978).

First, our lab tried to tackle information use as revealing each process, and soon afterward we turned to motivational moderators with more consistent success. Our earliest information-analytic work appeared only in book chapters (Fiske, 1982; Fiske and Pavelchak, 1986), as we struggled to find paradigms that would adequately test the CM informational predictions. Several studies manipulated category fit, showing that matches (i.e., category-consistent attributes) elicited evaluations consistent with the overall category, in what we termed *schema-triggered affect*. Mismatches (i.e., a category label but inconsistent attributes) elicited evaluations more in line with the attributes. One early study tested an old-flame hypothesis, whereby we tailored stimulus people to match (or not) each person's past or current partner on traits, appearance, or both. However appealing the idea and promising the preliminary results, we had confounded positivity with match. Susan Andersen later examined significant-other transference much more elegantly (for a summary and theory, see Andersen and Chen, 2002). Another converging effort, to manipulate politician schemas (mostly negative), suffered related confounds.

Yet another, but more convincing, study contrasted local stereotypes of engineers and artists, respectively “vegetables” and “fruits,” on that campus. By using each stereotype's content as the other's control, we finally unconfounded valence from category fit, by comparing stereotypic engineers to: other engineers with artistic traits, artists with engineer traits, and stereotypic artists. Thus, holding information constant, schema matches elicited negative responses in line with these negative stereotypes, more than did mismatches. The astute reader will note that, although we had unconfounded valence from consistency, the inconsistent stimulus people still might have been more interesting or salient.

Luckily for the CM, the first peer-reviewed evidence for dual processes of impression formation (Fiske et al., 1987) was an eventual pair of experiments designed to demonstrate category-based processes using two converging operationalizations, versus attribute-based processes, again using two (other) converging operationalizations. The main CM-related prediction was: affective responses will reflect the category when categorization is easy, but they will reflect the attributes when categorization is difficult.

The two easily categorized conditions provided either (1) a category label, followed by stereotype-consistent attributes (e.g., a loan shark who is opportunistic, shady, greedy, shrewd, and heartless); or (2) a category label, followed by uninformative attributes (e.g., an artist who is adult, medium height, employed, television-viewer, brown-haired). Two difficult-categorization conditions included either: (3) a category label, followed by category-inconsistent attributes (although consistent with another stereotype [p. 273 ↓] in the stimulus set; that is, a doctor who is bored, obedient, unenterprising, uneducated, and efficient does not fit a doctor stereotype, but does fit a maid stereotype); or (4) no category label, followed by attributes consistent with another stereotype in the stimulus set (e.g., a person who is practical, educated, scientific, skilled, and observant, attributes stereotypic for a doctor, but not immediately cuing the label). As a manipulation check, typicality ratings confirmed the predicted ease of categorization in each condition.

In a pretest, as [Table 13.1](#) indicates, participants had provided separate ratings of the category labels and the trait sets. This enabled us to calculate precise correlations between each person's likability rating for a particular stimulus combination and that same person's own prior separate ratings of the component category and attribute set, a technique originally crafted in my advisee Mark Pavelchak's dissertation (1989, see below). The critical correlations revealed that, in the two easily categorized conditions, a stimulus person's overall likability ratings correlated with the category labels (as well as the redundant or uninformative attributes). But in the two difficult-to-categorize conditions, likability ratings mainly correlated with the attributes but not the provided category label (which did not fit the attributes). What's more, in a second experiment, participants thought aloud into a tape recorder as they formed impressions. Many of the CM's hypothesized processes occurred: spontaneous comments about typicality, relying on the category when the attributes matched, recategorizing when they did not, and simply combining the attributes when no category fit.

Table 13.1 Stimulus materials for Experiments 1 and 2 in Fiske et al. (1987): categories, attributes, and their respective likabilities

Experiment 1				
<i>Loan shark</i>	<i>Artist</i>	<i>Doctor</i>	<i>Hotel maid</i>	<i>Person</i>
2.40	5.91	6.09	4.99	6.03
Opportunistic	Nonconforming	Practical	Bored	Adult
Shady	Creative	Educated	Obedient	Medium height
Greedy	Eccentric	Scientific	Unenterprising	Employed
Shrewd	Idealistic	Skilled	Uneducated	Television viewer
Heartless	Fashionable	Observant	Efficient	Brown-haired
2.05	6.23	6.46	3.91	5.38
Experiment 2				
<i>Professor</i>	<i>Reporter</i>	<i>Construction Worker</i>	<i>Politician</i>	<i>Person</i>
5.7	4.7	5.2	4.7	5.3
Intellectual	Curious	Strong	Selfish	Ordinary
Productive	Energetic	Loud	Power-hungry	Normal
Preoccupied	Perceptive	Rowdy	Pragmatic	Nice
Self-regulated	Aggressive	Red-neck	Opinionated	Usual
Hard-working	Liberal	Closed-minded	Smiley	Unremarkable
6.7	6.9	2.0	2.9	5.4

Note. All ratings are on 9-point scales, where 1 = not at all likable, 9 = extremely likable. Number of judges for Experiment 1 stimuli is 173; number of judges for Experiment 2 stimuli is 10, so fewer decimal places are given.

Simultaneous with this information-based mechanism for manipulating category-based versus attribute-based processes, Pavelchak (1989) developed the idea of manipulating the two processing modes by instruction. First, he developed trait lists that fit campus [p. 274 ↓] stereotypes for various academic majors (but did not automatically cue them). In an initial session, participants rated the likability of 35 majors and 50 relevant traits. In the second session, participants saw six stimulus people consisting only of four traits each. Half the participants first guessed the person's academic major (thereby activating a category) and then evaluated the person; the other half evaluated each person and then afterward guessed the academic majors. In the category-first condition, evaluations correlated with each participant's own prior evaluation of the category, more than with the average of the relevant attributes. In the no-category (piecemeal) condition, evaluations correlated with the average of the attributes, rather than the category that the participant guessed afterward. These (and some other Fiske-lab studies) laid the ground for informational conditions that moderate category and attribute-based processes.

At the same time, motivation shared our research agenda and soon dominated our work, for several reasons. First, my empirical talents emerged as stronger in interpersonal experimental social psychology than in heavily cognitive social psychology, as scores of old, unpublished manuscript files would prove. Second, on a more intellectual note, early on, Bill Swann and I met at Nag's Head workshops on social cognition, discovering that we shared a suspicion about the relentless person-

memory findings that people's free-recall favors expectancy-inconsistent information (e.g., Srull, 1981). This inconsistency bias simply could not hold in the hurly-burly of the real world, where people habitually rely on stereotypes and expectancies to bolster their memories; both our work attested to the power of category-based consistencies (Snyder and Swann, 1978; Taylor et al., 1978), and as later meta-analyses indeed indicated (Stangor and McMillan, 1992). How could the (stereotype) consistency and the (memory) inconsistency biases both be true? Maybe in a person-memory experiment with massive numbers of stimulus people (as in my own earlier work; Dreben et al., 1979; Fiske, 1980), perceivers do engage piecemeal processing that would favor inconsistencies, but otherwise, most of the time, in keeping with the CM, they should use their easily-at-hand category-based expectancies.

Swann and I wondered if it would matter whether people really cared about the other person; although we intended to run and later combine parallel collaborative studies, his interests shifted away. As it happens, my lab did design a study that showed no motivational effects on memory, but instead strong motivational effects on attention, attributions, and inferences. That study and its replications, heavily vetted by my local colleagues (as the acknowledgments to them and Swann attest), was our best paper to date, and probably my only empirical paper ever accepted outright, without revision (Erber and Fiske, 1984).

Graduate student Ralph Erber and I designed a situation in which people would expect to interact with another person, but in some cases would care, and in other cases not care, to go beyond superficial (category-based) impressions. We predicted that under ordinary circumstances, people would focus on expectancy-consistent information, but that if people cared about the person as a unique individual, they would go beyond their expectancies to focus on expectancy-inconsistent information. Although designed with stereotyping in mind, the experiment started with a simple positive-negative valenced expectancy.

Volunteers showed up to work as the novice in an expert-novice collaboration with a teacher-in-training (the expert). Together, they would design educational games for children, using silly wind-up toys. Undergraduates met a friendly but bland female confederate, who then exited, while they separately completed some introductory questionnaires. The partner's questionnaire communicated the valenced expectancy:

she predicted that she might be really good at this task, or that she thought her ability would be low. Apparently at our request, she had also [p. 275 ↓] brought with her some teaching comments allegedly written by her peers, each one a single sentence in different handwriting on a separate index card (this was the low-tech early 1980s). The comments pretested as positive were written in blue ink, the negative ones in black ink (or vice versa, across conditions). With a stopwatch in each jacket pocket (again, primitive but serviceable), the experimenter could surreptitiously time how long the participant examined each type of comment (blind to the match between color and valence). After reading the comments at their own speed to form a preliminary impression of their partner, participants rated her before an expected interaction that never came.

The crucial manipulation was motivation in the form of interdependence: how much they depended on their partner. Everyone expected to brainstorm ideas alone at first and then to work together with the partner. Half the participants expected to be eligible for a prize based on their work alone in the first half, while the other half expected to be eligible for the prize based on their joint work. We reasoned that the outcome-dependent participants would be motivated to pay more attention to their partner, especially to the most diagnostic information, the parts that disputed their expectancy, because they cared having an accurate understanding, to work together effectively. Indeed, outcome dependency uniquely predicted attention to inconsistent information (regardless of valence). Apparently in the service of deeper understanding, outcome-dependency also predicted more dispositional attributions, characterized as individuating the partner, beyond the simplistic expectancy. As the CM predicts, motivation moderated category- and attribute-based impression processes.

Conceptual replications of these studies bolstered their support for the CM. Actual stereotypes (of a recently discharged hospital patient, who was either an anxiety-provoking schizophrenic or a harmless heart patient) showed patterns similar to merely valenced expectancies, as well as effects on category-based versus attribute-based evaluations (Neuberg and Fiske, 1987). The crucial motivation appeared as an effort toward accuracy, consistent with the idea that people are trying to control their own outcomes by having a sense of accuracy about their partners who influence their potential success. Consistent with the accuracy mediator, people's motivations to use

category-based versus individuating processes respond to direct instruction (“try to be accurate”; Neuberg, 1989)

Interdependence presupposes correlated outcomes. If one's contingency on another person constitutes the crucial motivation, then – besides collaborative interdependence (positively correlated outcomes) – negative interdependence (competition) should show the same patterns. Interpersonal competition does exactly this (Ruscher and Fiske, 1990). More broadly, people respond to situational motivators in general. For example, people use the two processes according to perceived norms, for those who care about norms, and according to self-concept, for those who care about that (Fiske and Von Hendy, 1992). We were encouraged in all these studies by our colleagues, by then at University of Massachusetts, Amherst, and by National Science Foundation and National Institute of Mental Health support.

Having explored various moderators (motivation, instructions, norms, self-concept) and mediators (attention) predicted by the CM, the research program's next step was to explore boundary conditions. All the studies to date had concerned interpersonal interactions, where the correlated outcomes occurred within dyads; intergroup interdependence, having more complex contingencies, should not show the same effects (Ruscher et al., 1991). People prioritized individuating their teammates, with whom they had positive interdependence, but neglected the opposing team. Presumably, knowing the teammates is a necessary condition for understanding the group-on-group competition. Besides, the opponents can be assumed to have a simple shared purpose, [p. 276 ↓] namely besting one's own team. Learning their individual weaknesses can wait.

We further pursued the intergroup dynamics because an aggregate can operate as a monolithic, impenetrable group or as a bunch of knowable individuals. If one's outcomes depend on a hostile and homogenous group, presumably less open to influence, one might stay with categorical processes, but if the group were heterogeneous, presumably less tightly knit and therefore more open to influence, one might be motivated to use individuating processes. As predicted, the interpersonal differed from the intergroup context, apparently because of differences in the potential for control (Dépret and Fiske, 1999).

The perceiver's possibility of control began to appear more and more crucial. In an earlier study (Ruscher and Fiske, 1990), unconfident participants, who did not expect to do well at the joint task, simply gave up, not showing the same selective attention and dispositional inferences to the most diagnostic (inconsistent) information about their partners. The more we thought about interdependence as motivating attempts at accuracy, to control one's outcomes, the more explanatory value the concept of control seemed to have. A moderating role for control had also appeared in perceivers being subject to higher degrees of a homogenous outgroup's power (Dépret and Fiske, 1999), being a minority (Guinote et al., 2006), or being on power-holders' territory (Guinote and Fiske, 2003).

In other studies, perceived control also seemed to characterize our standard task-related outcome-dependency conditions, in which working harder or smarter at a specific task could presumably yield some control over one's outcomes. In contrast, however, under evaluation-related outcome dependency conditions, an absent powerholder would judge one's performance, without any interaction or personal contact, as is all-too-often the case. In this context, participants apparently felt they could not exert much control over their fate; instead they showed a positivity bias, hoping for the best from their evaluator (Stevens and Fiske, 2000). Similarly, in another setting that entailed an evaluation of oneself as an overall person – romantic outcome-dependency – participants again undermined accuracy motivation, settling for wishful thinking (Goodwin et al., 2002). When unable to control their fate, people apparently hope for the best.

During the course of these studies that established for us the importance of control, another variable also yielded a more precise understanding as we operationalized it in different ways. Interdependence implies that two people each depend on the other, as in one-on-one cooperation or competition. So far, we had shown that mutual interdependence increases motivations to be accurate and control one's outcomes, resulting in attention to diagnostic information, dispositional inferences that individuate the other person, and relatively attribute-based impressions. Theoretically, though, we had a confound. The interdependence could cause people either (1) to worry about their own, contingent outcomes (our hypothesis) or (2) to feel responsible for the other person; either case could motivate people to form more attribute-based, careful impressions. Two lines of work resulted from this insight.

First, a series of studies isolated the participant's own outcome dependency on (versus independence from) the partner. For example, people who depend on their partner, either symmetrically or asymmetrically, respond in similar ways. Compared with an independent control group, asymmetrically outcome-dependent participants – that is, those who depend on their partner but the partner does not depend on them – still strive for accuracy and control (Stevens and Fiske, 2000).

Second, we began to take power seriously. So far, we had examined outcome dependency from the bottom up. What about the top-down perspective? Defined as control over another's outcomes, power should make people neglect the other person as a unique individual, according to the CM logic, because the powerful are less motivated to go [p. 277 ↓] beyond their default categorical impressions. A theory of power as control predicted: The powerful “need not, cannot, and want not” to attend to their subordinates because of, respectively, their own outcome control, the sheer numbers of subordinates, and potential personality correlates of seeking power (Fiske, 1993: 621). Laying out these predictions did not immediately recruit graduate students to collaborate on the project. American students seemed to react as if power were a rude, taboo topic. Who wants to study Machiavellian people who manipulate other people? A French advisee, however, was the first enthusiast, and power as asymmetrical control informed our studies of power from the bottom-up (Dépret and Fiske, 1999, as noted).

Then, finally, some other graduate students wanted to examine the powerholders themselves; that is, people vulnerable to stereotyping others whom they do not need for anything. Powerholders with more control indeed engage in more stereotyping than powerholders with less control, and trait dominance has similar effects (Goodwin et al., 2000). Individually dominant power-holders do not necessarily attune to the task-relevant competence of their subordinates, whereas low-dominant powerholders do (Operario and Fiske, 2001). Consistent with being clueless, the dominant powerholders are entertained and pleased by sociable subordinates, but they do not retain them for the task at hand.

Nearly a decade after the original published theory (Fiske and Neuberg, 1990) and almost two decades after the first published chapter (Fiske, 1982), we re-evaluated the CM (Fiske et al., 1999). Evidence for the model suggested that its premises had aged

gracefully. Several motives moderate category-based and individuating processes, parsed as belonging, understanding, controlling, self-enhancing, and trusting (Stevens and Fiske, 1995). Various individual differences also predispose people toward one kind of process or the other. For example, need for structure encourages category use (Neuberg and Newsom, 1993), whereas need for cognition encourages attribute-oriented processing (Cacioppo et al., 1996).

Two primary debates remained, both oriented to fundamental cognitive processes. One debate contrasted the CM's apparently serial processes (first categorization, then attribute-based processes) with parallel, simultaneous processes together constraining outcomes (Kunda and Thagard, 1996). The original CM was agnostic about the underlying deep-cognitive processes. CM does claim that category processing typically will prove faster, but attribute processing could co-occur, just more slowly, so that its results appear later. Or category and attribute processing could occur serially, the latter only as needed. Most likely, because attributes and categories influence each other, parallel processes make sense. However, this deep cognitive modeling was not the CM's main level of analysis (see Fiske et al., 1999, for more detail).

Another debate centered on the nature of the representations for category-based and attribute-based impressions. A contemporaneous dual-process impression model (Brewer, 1988), like the CM, posited two impression-formation processes, one more stereotypic, and the other more personal. Several factors differentiate the two models (Fiske, 1988), but the primary differences lie in Brewer's model of branching decision-tree process, rather than a continuum, and in its proposal that distinct branches represent people in distinct forms (category, image, exemplar). Other differences include the CM's explicit focus on attention and interpretation as mediators, motivation as a moderator, a common set of decision rules (information fit), and a common representational format. Brewer's model, which posits branches rather than a continuum, proposes not only distinct representations after each decision point, but also distinct decision rules at each decision point (for more detail, see Brewer and Harasty Feinstein, 1999).

Two-mode models evidently ruled the turn of the century, as witnessed by the array of three dozen such models, too numerous to [p. 278 ↓] review here but collected in an edited volume (Chaiken and Trope, 1999). The larger conceptual point is that

the contrast between relatively automatic, rapid judgments (“System 1”) and relatively thoughtful, slower judgments (“System 2”) has permeated social cognition, attitudes, and more (see Fiske and Taylor, 2008, chapter 2, for an overview). Like other dual-process models, the CM does unify apparently separate processes, specifies conditions motivating each, and describes processes underlying both. The added value is moving from either-or to both-and.

The related theory of power-as-control generated its own controversies. For example, we had asserted (and shown) that power-holders, because of their noncontingency on others, are vulnerable to category-based processing, neglecting their subordinates (Goodwin et al., 2000). Just as our oft-presented power results got into print, other studies showed the boundaries of our predictions. For example, powerholders oriented toward social responsibility do individuate their subordinates (Overbeck and Park, 2001). Exchange-oriented powerholders stereotype, but communally oriented powerholders do not (Chen et al., 2001). The power literature then exploded with the theory of power as an approach-driven motivation. Consistent with our view but independent and much more comprehensive, a far-reaching theory showed that the powerful see the world as full of rewards, and powerholders cheerfully operate on automatic (Keltner et al., 2003). What's more, powerholders respond more constructively to what the situation affords them (Guinote, 2008), and they act on their goals (Galinsky et al., 2003). The review of this new power literature lies beyond the current chapter's scope (see Fiske, 2010), but some aspects trace back to the CM-based power-as-control theory.

Much remains to do. Future CM challenges include pursuing neural correlates of the two types of process (Lieberman et al., 2002). Early evidence indicated separate neuropsychological processes involved in generalizations (category use) and elaboration (attribute use), differentially vulnerable to cognitive decline with aging (Mather et al., 1999). Recently we observed undergraduates' dual neural processes in the scanner, while they formed impressions of two partners, when outcome-dependent on one but not the other (Ames and Fiske, 2011). As before, they concentrated on the most informative, expectancy-inconsistent cues, but only for the partner on whom their outcomes depended. In this case only, the medial prefrontal cortex selectively activated—in an area independently identified as responsive to impression formation

processes. Future research will further differentiate neural systems associated with the two kinds of impression processes.

On another promising note, in provocative early social neuroscience studies, novel other-race faces activated the amygdala, a brain region implicated in emotional vigilance (Hart et al., 2000; Phelps et al., 2000). These early returns encouraged us to pursue category-based and more individuating impressions via neuroimaging methods. After I moved to Princeton, graduate student Elizabeth Wheeler and I manipulated category-based and more individuating responses to novel cross-race yearbook photos. Based on the CM, we noted that the default conditions used in prior studies (categorizing by gender, for example) would tend to encourage category-based processing. Instead, we compared three conditions (Wheeler and Fiske, 2005): the default, categorical processing (report the pictured person's gender); a totally nonsocial condition (does the photo show a dot on the face?); and a newly-invented individuating condition (would the pictured person like a just-pictured vegetable?). Consistent with the CM, the category condition replicated the race-based amygdala activation, but neither the nonsocial nor the individuating condition did so. A conceptual replication with another outgroup (homeless people) showed that the individuating condition uniquely activated the medial prefrontal cortex (mPFC), a region implicated in social attributions (Harris and Fiske, 2006).

[p. 279 ↓]

Describing Categories: The Stereotype Content Model

The CM privileges categories as capturing much impression formation, being the fastest, the default, and the anchoring process. If categories matter so much, which categories matter and why (Puzzle #7)? The stereotype content model (SCM) holds that, just as the processes of impression formation reflect systematic principles, so too do the contents of our categories. The SCM complements and extends the CM, in the intellectual narrative of theoretical development and empirical epic.

The SCM's fundamental insight is simple (for an overview, see Fiske et al., 2007; for detail, see Cuddy et al., 2008). The dimensions people use to categorize other people are predictable because the human problem is universal. People as perceptual objects mainly differ from nonhuman entities in having intention; that is, being autonomous agents (Fiske and Taylor, 2008: Chapter 1). So, other people need to know, first, whether another social entity (individual or group) intends the self good or ill. The sentry's query, "Friend or foe?" captures this dilemma. We call this dimension warmth (warm, friendly, trustworthy, honest). People judge this dimension of a face in less than a second (Willis and Todorov, 2006). The second question is whether the other can enact that intention: How able or unable? The competence dimension (competent, capable, skilled) also operates rapidly, only slightly more slowly than the first dimension.

Together, these two dimensions, under various names, account for 80 to 90 percent of the variance in individual impressions (Wojciszke, 2005). Solomon Asch (1946), again, was there first, but only intuitively, when his experiments held competence constant (*intelligent, skillful, industrious, determined, practical, cautious*) but contrasted warmth (*warm* versus *cold*). Fritz Heider's (1958) astute phenomenological analysis, as usual, identified the relevant factors; he described people inferring dispositions by understanding what another person may *try* to do (goal, which includes warm or hostile intent) versus what the other *can* do (ability). On a similar but more empirical note, Seymour Rosenberg (Rosenberg and Sedlak, 1972) factor analyzed personality trait impressions, generating two dimensions: social good-bad (anchored by *warm* or *sociable* versus *unpopular*) and task good-bad (anchored by *foolish* versus *scientific*). Reflecting their breadth, similar dimensions of competence and integrity constitute impressions of presidential candidates (Abelson et al., 1982).

SCM Hypothesis 1 states that the two dimensions will differentiate societal groups into predictable quadrants. Indeed, common societal groups do spread out across the dimensions in the U.S. (Fiske et al., 2002; Cuddy et al., 2007), and all over the world (Cuddy et al., 2009). As Table 13.2 indicates, the dimensions combine to define universally recognizable stereotypes. The high-warmth/high-competence groups typically represent the ingroup and its allies, as well as societal reference groups, such as the middle class, heterosexuals, the dominant religion (e.g., in the US, Christianity), and the majority racial group. The outgroup quadrants include the lowest of the low, seen as neither warm nor competent: poor people all over the world, and in the US

specifically, homeless people and drug addicts. They allegedly have no redeeming features. This much (high/high ingroups versus low/low outgroups) overlaps with standard depictions of intergroup relationships.

	Low competence stereotypes from low status	High competence stereotypes from high status
High warmth stereotypes from high cooperation		
Example	Old people	Middle-class people
Affect	Pity	Pride
Behavior (active ... and passive)	Active help	Active help
	Passive neglect	Passive accommodation
Low warmth stereotypes from low cooperation		
Example	Homeless people	Rich people
Affect	Disgust	Envy
Behavior (active ... and passive)	Active harm	Active harm
	Passive neglect	Passive accommodation

The SCM contributes the novel idea of mixed, ambivalent stereotypes, high on one dimension and low on the other. Some ambivalent outgroup quadrants include those high-warmth, low-competence groups (e.g., older people, disabled people), that are benign, harmless, but do not especially bear on ingroup goals. They may be likable, but not especially respected. Other ambivalent outgroups include those low-warmth, high-competence groups (e.g., rich people, outsider entrepreneurs), who earn grudging respect but also disliking.

SCM Hypothesis 2 states that many or even most groups will locate in mixed quadrants. Across the world, the majority of societal groups do indeed fall into the two ambivalent clusters (Cuddy et al., 2008, 2009). However, a society's degree of ambivalence depends on its degree of inequality: the larger its income gap, the more the society apparently has to justify the divisions between its groups, so more of them are favored on one dimension (e.g., rich people may be competent) but not the other (rich people are cold) (Durante et al., submitted).

The SCM posits that stereotype contents are systematic not only by basic dimensions but also by their antecedents and consequences: *Social structure causes perceived stereotypes, which in turn generate affect, which in its turn generates behavior.* Having described the stereotype contents, what are their antecedents? The stereotypes' origins prove predictable. Social structure predicts perceived traits, based on two principles.

SCM Hypothesis 3 states that perceived competition lowers perceived warmth, and perceived status raises perceived competence. SCM predicts the competition-warmth link from a previous analysis of why people feel negative toward outgroups (Fiske and Ruscher, 1993). Negative affect toward most outgroups follows from the inherent goal conflict: the group is an outgroup precisely because their goals do not support or may even conflict with ingroup goals. Hence, they cannot be trusted, seen as having malignant or at least not benign intent. The competition-warmth link averages a small but reliable effect across American and European samples (Cuddy et al., 2008). With improved measurement of both competition (to include value conflict as well as resource conflict) and warmth (to include not only sociability but also trustworthiness), consistent with the SCM, the correlations prove substantial (Kervyn et al., submitted).

On the other dimension, the status-competence link follows from dispositional attributions for a high-status position being due to the ability of the incumbents. This also follows from a just-world belief that groups get what they deserve, and indeed those higher on belief in a just world show even stronger status-competence correlations than people lower on this dimension (Oldmeadow and Fiske, 2007). Around the world, the status-competence correlation is reliable and averages large in size (Cuddy et al., 2008).

Turning to the consequences of stereotype content, according to *SCM hypothesis 4*, each cluster provokes predictable emotional prejudices (see [Table 13.2](#)). In the US and Europe, the supposedly both warm and competent ingroup and its allies typically elicit pride because they are responsible for positive outcomes that reflect well on the self. The low-warmth/low-competence groups are allegedly responsible for their own negative outcomes; they elicit disgust, an unadulterated negative [p. 281 ↓] emotion applicable to contaminated objects as well as people.

The warm but incompetent outgroups (e.g., the elderly) elicit pity, because they have negative outcomes that are not their fault. In contrast, the competent but cold outgroups (e.g., rich people) elicit envy because they have positive outcomes that do not favor the ingroup. Indeed, the ingroup may resent them for their good fortune that relatively deprives the ingroup. Both pity and envy are mixed emotions, so these two quadrants qualify as ambivalent on two counts: stereotype content that is high on one fundamental dimension and low on the other, plus mixed emotions.

Besides social structure predicting correspondent stereotypes and both predicting unique emotional prejudices, *SCM Hypothesis 5 predicts behavioral tendencies distinct to each quadrant* (see [Table 13.2](#); Cuddy et al., 2007). We hypothesized that warmth, being primary, would predict active responses to facilitate or interfere with those respectively allied or opposed to ingroup interests. Thus, high-warmth groups obtain active help and protection; low-warmth groups elicit active attack and fight. Both ingroup and pitied outgroups receive help and protection, for example, because both are stereotyped as warm.

The competence dimension, being secondary, should elicit passive facilitation (for competent groups) and passive interference (for incompetent groups). Thus, both ingroup and envied outgroups, both being stereotypically competent, elicit passive association, going-along-to-get-along, while disgusting outgroups and pitied outgroups, both being stereotypically incompetent, receive neglect and exclusion. A national random sample survey (Cuddy et al., 2007) supports all these predictions. For example, the ambivalently stereotyped pitied outgroups receive both active protection and passive neglect (e.g., institutionalizing the disabled or elderly). Another ambivalently stereotyped, but instead envied, cluster receives both passive association (e.g., shopping at their stores) but also active harm, at least under some circumstances (e.g.,

looting and attack, under social breakdown). Across quadrants, emotions are proximate cause of behavior.

Sorting groups along these two dimensions, warmth and competence, arguably begins to look universal. If so, the historical stereotypes should show similar patterns, and they do. We replicated the SCM using recoded Princeton studies of ethnic and national stereotypes since the 1930s (Bergsieker et al., submitted). Furthermore, content coding of fascist magazine descriptions of ethnic and national groups also fits the SCM space (Durante et al., 2009).

Some national and historical variations do occur. For example, the high-high quadrant, normally reserved for the ingroup and its allies, is relatively vacant in three East Asian samples (Cuddy et al., 2009). Instead, various ingroups migrate to a more modest, middling position, in keeping with cultural norms for modesty. And among the fascists, the pity quadrant is empty, a finding that speaks for itself.

The SCM subsumes and explains particular ambivalent prejudices. For example, Peter Glick and I had previously hypothesized that sexism is not simple antipathy but instead ambivalent (Glick and Fiske, 1996). Although male status is higher in all societies, men and women universally depend on each other to procreate, as well as to form lasting bonds. This combined status difference and obligatory cooperation generates a default stereotype of traditional women, termed benevolent sexism (paternalism). The more expected form, hostile sexism, targets nontraditional women as admittedly competent but cold (e.g., lesbians, feminists, career women). Although ambivalent sexism theory (AST) predates the SCM, AST emerged from a social structural analysis of gender prejudice that anticipated the SCM's focus on cooperation-competition and societal status.

Another source of inspiration for AST were the Katz and Hass (1988) studies of racial ambivalence, which inspired us to consider the possibility that sexism too could be ambivalent, though for different reasons. [p. 282 ↓] Regardless of its origins, the AST patterns occur worldwide, consistent with their origins in the fundamental relationships between men and women (Glick et al., 2000). In SCM terms, traditional women often inhabit the high-warmth, low-competence (pitied) quadrant, while nontraditional women inhabit the low-warmth, high-competence envied quadrant.

Other groups occupy combinations of quadrants, differentiated by subtypes. For example, stereotypes of black Americans differentiate by social class into black professionals, who make others proud, versus poor blacks, who disgust others (Fiske et al., 2009). Stereotypes of gay men fall into all quadrants, depending on subtype (e.g., in-the-closet, gay artist, leather-biker, cross-dresser; Clausell and Fiske, 2005). And although the default stereotype of old people is high-warmth, low-competence (Cuddy et al., 2005), subtypes of older people may try to escape, forfeiting protective pity for envious resentment directed at those who hoard their wealth, disgust at those who selfishly consume societal resources, or resistance toward those who attempt to invade younger ingroup identities (North and Fiske, submitted). Particular prejudices thus come into focus, viewed through the SCM lens. For example, anti-Asian prejudice fits the patterns of the envy quadrant (Lin et al., 2005). Dehumanization of homeless people fits the patterns of the disgust quadrant (Harris and Fiske, 2006). Envy of investment bankers and schadenfreude at their misfortunes fits the pattern of the envy quadrant (Cikara and Fiske, submitted). Likewise, the fans of mutually envied sports rivals (e.g., Yankees versus Red Sox) experience schadenfreude at their rivals losses, even to a third party (Cikara et al., 2011). Altogether, the data include self-report questionnaires, individual differences, neural patterns, and subtle facial expressions indicated by electromyography.

The original SCM research initially assessed perceived societal responses to various groups, leaving it open to criticism that it reflected merely ideology. And all the evidence was correlational, raising questions about the proposed causal relationships. Subsequent work revealed the SCM to be both personal and causal. Experimental vignette studies showed that manipulating the social structures of interdependence and status does result in the predicted warmth and competence stereotypes (Caprariello et al., 2009). Laboratory analog experiments manipulating interpersonal interdependence and status showed that the structure of dyadic relationships also predicts personal warmth and competence expectancies (Russell and Fiske, 2008). Finally, people report personal emotional reactions to images of groups from the different quadrants, in line with SCM predictions (Harris and Fiske, 2006).

Neuroimaging data show emotional reactions consistent with self-reports so far for two quadrants, the first being disgust toward drug addicts and homeless people; insula activation typically characterizes disgust and in our data insula differentially activates

to those images (Harris and Fiske, 2006). Second, the schadenfreude toward envied sports rivals appears in ventral striatum reward-area activation that correlates with self-reported harm to their fans (Cikara et al., 2011).

Future research will focus on patterns related to specific social groups, such as higher and lower classes, especially interpersonal effects of social structure, such as trust, entitlement, and imposter feelings. Where possible, we will pursue psycho-physiological as well as self-report indicators.

Application to Social Issues

Both the CM and the SCM instantly bolted from the ivory tower to risk real-world application. The CM particularly proved useful to discrimination lawsuits, in which I served as an expert witness. In testimony, the CM helped explain the *social framework* of bias in organizations (Fiske and Borgida, 2008), particularly how decision makers' relatively [p. 283 ↓] automatic categories could lead them to discriminate in cool, cognitive, categorical ways, without any emotional animus:

- A female welder complained about workplace pornography that encouraged sexual harassment of the few women employees (*Robinson v. Jacksonville Shipyards, Inc.*¹). The prominently displayed sexualized media both primed objectified categories for women and trivialized their distress when they reported the resulting problems.
- Female applicants for jobs at a home-improvement franchise were routinely routed to deadend cashier jobs, while comparably qualified male applicants were sent to the sales floor, on their way to bonuses and promotion (*Butler et al. and Frank et al. v. The Home Depot Inc.*²). Interviewers categorized most men and few women as competent in construction, undervaluing the self-selected female candidates' knowledge of home renovation,
- Evaluators demanded that a female manager behave like a more stereotypic woman (*Hopkins v. Price Waterhouse*³). Although the top earner and expert in her cohort, she was advised to walk, talk, and dress in more feminine ways if she wanted to make partner.

The CM proves useful because lay people (including judges and juries) do not expect that categories, without active antipathy, could cause discrimination. For example, accounting executive Ann Hopkins was a top producer at Price Waterhouse, but she was denied partnership because of alleged interpersonal skills deficits that stemmed from being insufficiently feminine. Essentially, she made people uncomfortable because she did not fit her stereotypic category. Psychological science, including the CM, helped frame judges' understanding, all the way to the Supreme Court (Fiske et al., 1991).

The testimony in turn affected the theories. In preparing expert reports, I noticed a misfit between the available social psychological literature, which studies peers' stereotypic judgments, and real-world organizations, in which powerful decision makers' stereotypic judgments matter most. The power-as-control theory (Fiske, 1993) directly resulted from expert testimony experiences, in interplay with the CM interdependence hypotheses. Asymmetrical control over outcomes (power) should make people vulnerable to stereotyping their subordinates, as noted earlier.

The SCM also has applied to several social issues. The potential for people to dehumanize some types of outgroups informed a *Science Policy Forum* article about the psychological dynamics of prisoner abuse at Abu Ghraib (Fiske et al., 2004). The SCM-related ambivalent sexism theory has informed expert testimony by several social psychologists (Rudman et al., 2008). And the SCM-predicted role of emotions in mediating the impact of stereotypes on behavior has informed a meta-analysis of how best to predict racial discrimination (Talaska et al., 2008).

Relevant to all the theories described here, common sense underestimates the automaticity, ambivalence, and ambiguity of prejudice. The legal implications argue against judges and juries relying on their lay psychological theories, for example, that only conscious animus causes discrimination. A *behavioral realism* approach argues for the relevance of established psychological science in the courtroom, for example, in employment law (Krieger and Fiske, 2006). Psychological theories, supported by evidence, can inform policies to avert discrimination, by suggesting organizational remedies, such as transparency about an organization's success in promoting underrepresented groups (Fiske and Krieger, in press).

Social cognition theories, such as the CM and SCM, provide expertise in employment cases, particularly drawing on the high-quality science establishing (1) dual processes (automatic and controlled), (2) early bias (at attention and encoding), and (3) mental construal (interpretation) that creates category-based perception (Fiske and Borgida, 2008). The relevance of the CM is clear to all three applications, but also, respectively, are the Implicit Association Test regarding automaticity (e.g., Kang and Banaji, 2006), the [p. 284 ↓] social neuroscience of prejudice at the earliest stages of encoding (e.g., Phelps et al., 2000), and the tenacity of categories despite individuating information (e.g., Heilman and Haynes, 2008). Further, a variety of psychological science usefully informs legal practice and policy, beyond common sense (Borgida and Fiske, 2008).

Conclusion

According to the opening personal narrative, the two theories of social category process and content – CM and SCM – have roots in the puzzles of lived experience. People are captured by categories, and consequently my work has been captivated by categories. Many theories explain the puzzles narrated earlier, but here, consider how CM and SCM explain everyday life.

Why did my by-chance dance partner in high school need the additional category of knowing my astrological sign? Because, as the CM predicts, he was motivated to go beyond gender, age, and race, and the new information evidently recategorized his impression (Puzzle #1 solved). Why did all the kids sort themselves by ethnicity in the middle-school cafeteria? Because, as the CM predicts, they were relying on easy categories, and as the SCM predicts, their ingroup elicited pride, thereby avoiding pity, envy, or disgust they might feel toward other cliques (Puzzle #2 solved).

Why are some communities, such as Hyde Park, able to integrate, while others remain segregated? At the social psychological level, part of the success is predicted by the contact hypothesis (Allport, 1954), which names interdependence as a key condition. The CM shows that cooperation makes people attend to each others' counter-stereotypic features, enabling them to overcome their categories and work together toward common goals (Puzzle #3 solved). The CM also identifies personal values, certainly relevant to that community that prizes its multicultural character.

Why did the college doorman think I was Kathleen Kennedy, based only on demographics? As the CM predicts, he was not motivated to go beyond superficial categories. And maybe some wishful thinking entered in. Puzzle #4 yields to the primacy of visual demographic categories, as predicted by the CM.

Why are likability and memory for the original evidence so separate, as in my ability to recall my first students' likability more than their identity? As the CM predicts, people form online impressions resulting in affective tags that persist beyond the original inferences. Puzzle #5 solved.

And finally, regarding the CM, how can one reconcile both category-based and attribute-based processes, as the Ohio State graduate student's question that generated Puzzle #6? Motivation and information both decide which process operates when.

As for the SCM, Puzzle #7 asked whether stereotype contents are systematic, and indeed, they are, along two fundamental dimensions of warmth and competence, with social structural predictors, associated affect, and tendencies toward behavior.

At this point, several puzzles have yielded, and some theories have developed with collaborators who inspire equal doses of curiosity, insight, persistence, and delight. Together, we faced many obstacles and dead ends, but building theory is not a solo sprint, it takes a marathon team. And that's our favorite category.

Acknowledgments

The author acknowledges Russell Sage Foundation for support during the writing of this chapter and Lydia Emery for comments on a prior draft.

Notes

1 *Robinson v. Jacksonville Shipyards, Inc.* [M.D. Fla. 1989]; Case No. 86-927.

[p. 285 ↓] 2 *Butler et al. and Frank et al. v. The Home Depot Inc.* [1994, 1995]. US District Court, N. District of California, C 94–4335 SI and C 95–2182 SI.

3 *Hopkins v. Price Waterhouse*, 618 F. Supp. 1109 [D.D.C. 1985]; appeal: *Price Waterhouse v. Hopkins*, 825 F.2d 458 [D.C. Cir. 1987]; Supreme Court review: *Price Waterhouse v. Hopkins*, 109 S. Ct. 1775 [1989]; remand: *Hopkins v. Price Waterhouse*, No. 84–3040, slip op. [D.D.C. May 14, 1990].

References

Abelson, R.P., Kinder, D.R., Peters, M.D. and Fiske, S.T. (1982) Affective and semantic components in political person perception . *Journal of Personality and Social Psychology* , 42, 619–630. <http://dx.doi.org/10.1037/0022-3514.42.4.619>

Allport, G.W. (1954) *The Nature of Prejudice* . Reading, MA: Addison-Wesley.

Ames, D.L. and Fiske, S.T. (2011) Outcome dependency alters the neural substrates of impression formation . Unpublished manuscript.

Andersen, S.M. and Chen, S. (2002) The relational self: An interpersonal social-cognitive theory . *Psychological Review* , 109, 619–645. <http://dx.doi.org/10.1037/0033-295X.109.4.619>

Anderson, N.H. (1981) *Foundations of Information Integration Theory* . New York: Academic Press.

Asch, S.E. (1946) Forming impressions of personality . *Journal of Abnormal and Social Psychology* , 41, 1230–1240.

Bergsieker, H.B., Leslie, L.M., Constantine, V.S. and Fiske, S.T. (submitted) Stereotyping by omission: Eliminate the negative, accentuate the positive.

Borgida, E., ed. and Fiske, S.T. (eds) (2008) *Psychological Science in the Courtroom: Beyond Common Sense* . London: Blackwell.

Boring, E.G. (1950) A History of Experimental Psychology , 2nd Edition. Englewood Cliffs, NJ: Prentice-Hall.

Brewer, M.B. (1988) A dual process model of impression formation . In T.K. Srull, ed. and R.S. WyerJr. (eds), Advances in Social Cognition , 1, 1–36. Hillsdale, NJ: Erlbaum.

Brewer, M.B. and Harasty Feinstein, A.S. (1999) Dual processes in the cognitive representation of persons and social categories . In S. Chaiken, ed. and Y. Trope (eds), Dual-process Theories in Social Psychology , pp. 255–270. New York: Guilford Press.

Cacioppo, J.T., Petty, R.E., Feinstein, J.A. and Jarvis, W.B.G. (1996) Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition . Psychological Bulletin , 119, 197–253. <http://dx.doi.org/10.1037/0033-2909.119.2.197>

Cantor, N. and Mischel, W. (1977) Traits as prototypes: Effects on recognition memory . Journal of Personality and Social Psychology , 35, 38–48. <http://dx.doi.org/10.1037/0022-3514.35.1.38>

Caprariello, P.A., Cuddy, J.C. and Fiske, S.T. (2009) Social structure shapes cultural stereotypes and emotions: A causal test of the stereotype content model . Group Processes and Intergroup Behavior , 12, 147–155. <http://dx.doi.org/10.1177/1368430208101053>

Chaiken, S., ed. and Trope, Y. (eds) (1999) Dual-process Theories in Social Psychology . New York: Guilford Press.

Chen, S., Lee-Chai, A.Y. and Bargh, J.A. (2001) Relationship orientation as a moderator of the effects of social power . Journal of Personality and Social Psychology , 80, 173–187. <http://dx.doi.org/10.1037/0022-3514.80.2.173>

Cikara, M., Botvinick, M.M. and Fiske, S.T. (2011) Us versus them: Neural responses to rival groups' pleasures and pains . Psychological Science . 22, 306–313. <http://dx.doi.org/10.1177/0956797610397667>

Cikara, M. and Fiske, S.T. (submitted) Stereotypes and Schadenfreude: Behavioral and physiological markers of pleasure at others' misfortunes. Unpublished manuscript.

Clark, M.S., ed. and Fiske, S.T. (eds) (1982) Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition . Hillsdale, NJ: Erlbaum.

Clausell, E. and Fiske, S.T. (2005) When do the parts add up to the whole? Ambivalent stereotype content for gay male subgroups . Social Cognition , 23, 157–176. <http://dx.doi.org/10.1521/soco.23.2.161.65626>

Cuddy, A.J.C., Fiske, S.T. and Glick, P. (2007) The BIAS map: Behaviors from intergroup affect and stereotypes . Journal of Personality and Social Psychology , 92, 631–648. <http://dx.doi.org/10.1037/0022-3514.92.4.631>

Cuddy, A.J.C., Fiske, S.T. and Glick, P. (2008) Competence and warmth as universal trait dimensions of interpersonal and intergroup perception: The Stereotype Content Model and the BIAS Map . In M.P. Zanna (ed.), Advances in Experimental Social Psychology , 40, 61–149. New York: Academic Press.

Cuddy, A.J.C., Fiske, S.T., Kwan, V.S.Y., Glick, P., Demoulin, S., Leyens, J-Ph., Bond, M.H., Croizet, J.-C, Ellemers, N., Sleebos, E., Htun, T.T., Yamamoto, M., Kim, H.-J., Maio, G., Perry, J., Petkova, K., Todorov, V., Rodriguez-Bailón, R., Morales, E., Moya, M., Palacios, M., Smith, V., Perez, R., Vala, J. and Ziegler, R. (2009) Is the stereotype content model culture-bound? A cross-cultural comparison reveals systematic similarities and differences . British Journal of Social Psychology , 48, 1–33. <http://dx.doi.org/10.1348/014466608X314935>

Cuddy, A.J.C., Norton, M.I. and Fiske, S.T. (2005) This old stereotype: The pervasiveness and persistence of the elderly stereotype . Journal of Social Issues , 61, 265–283. <http://dx.doi.org/10.1111/j.1540-4560.2005.00405.x>

Dépret, E.F. and Fiske, S.T. (1999) Perceiving the powerful: Intriguing individuals versus threatening groups . Journal of Experimental Social Psychology , 35, 461–480.

Dreben, E.K., Fiske, S.T. and Hastie, R. (1979) Impression and recall order effects in behavior-based impression formation . *Journal of Personality and Social Psychology* , 37, 1758–1768. <http://dx.doi.org/10.1037/0022-3514.37.10.1758>

Durante, F., Fiske, S.T., Cuddy, A.J.C. and Kervyn, N. (submitted) Inequality predicts ambivalence in stereotype content. Unpublished manuscript.

Durante, F., Volpato, C. and Fiske, S.T. (2009) Using the Stereotype Content Model to examine group depictions in fascism: An archival approach . *European Journal of Social Psychology* , 39, 1–19. <http://dx.doi.org/10.1002/ejsp.599>

Erber, R. and Fiske, S.T. (1984) Outcome dependency and attention to inconsistent information . *Journal of Personality and Social Psychology* , 47, 709–726. <http://dx.doi.org/10.1037/0022-3514.47.4.709>

Fiske, S.T. (1980) Attention and weight in person perception: The impact of negative and extreme behavior . *Journal of Personality and Social Psychology* , 38, 889–906. <http://dx.doi.org/10.1037/0022-3514.38.6.889>

Fiske, S.T. (1982) Schema-triggered affect: Applications to social perception . In M.S. Clark, ed. and S.T. Fiske (eds), *Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition* , pp. 55–78. Hillsdale, NJ: Erlbaum.

Fiske, S.T. (1988) Compare and contrast: Brewer's dual-process model and Fiske et al.'s continuum model . In T.K. Srull, ed. and R.S. Wyer (eds), *Advances in Social Cognition*, Vol. 1: *A Dual Model of Impression Formation* , pp. 65–76. Hillsdale, NJ: Erlbaum.

Fiske, S.T. (1993) Controlling other people: The impact of power on stereotyping . *American Psychologist* , 45, 621–628. <http://dx.doi.org/10.1037/0003-066X.48.6.621>

Fiske, S.T. (2010) Stratification . In S.T. Fiske, ed. , D.T., Gilbert, ed. and G. Lindzey (eds), *Handbook of Social Psychology* , 5th Edition. New York: Wiley.

Fiske, S.T., Bergsieker, H., Russell, A.M. and Williams, L. (2009) Images of Black Americans: Then, 'them' and now, 'Obama!' *DuBois Review: Social Science Research on Race* , 6, 83–101. <http://dx.doi.org/10.1017/S1742058X0909002X>

Fiske, S.T., Bersoff, D.N., Borgida, E., Deaux, K. and Heilman, M.E. (1991) Social science research on trial: The use of sex stereotyping research in *Price Waterhouse v. Hopkins* . *American Psychologist* , 46, 1049–1060. <http://dx.doi.org/10.1037/0003-066X.46.10.1049>

Fiske, S.T. and Borgida, E. (2008) Providing expert knowledge in an adversarial context: Social cognitive science in employment discrimination cases . *Annual Review of Law and Social Science* , 4, 123–148. <http://dx.doi.org/10.1146/annurev.lawsocsci.4.110707.172350>

Fiske, S.T., Cuddy, A.J.C. and Glick, P. (2007) Universal dimensions of social perception: Warmth and competence . *Trends in Cognitive Science* , 11, 77–83. <http://dx.doi.org/10.1016/j.tics.2006.11.005>

Fiske, S.T., Cuddy, A.J.C., Glick, P. and Xu, J. (2002) A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition . *Journal of Personality and Social Psychology* , 82, 878–902. <http://dx.doi.org/10.1037/0022-3514.82.6.878>

Fiske, S.T., Harris, L.T. and Cuddy, A.J.C. (2004) Policy Forum: Why ordinary people torture enemy prisoners . *Science* , 306, 1482–1483. <http://dx.doi.org/10.1126/science.1103788>

Fiske, S.T. and Krieger, L.H. (in press) Policy implications of unexamined discrimination: Gender bias in employment as a case study. In E. Shafir (ed.) *Behavioral Foundations of Policy* . Princeton: Princeton University Press and New York: Russell Sage.

Fiske, S.T., Lin, M.H. and Neuberg, S.L. (1999) The Continuum Model: Ten years later . In S. Chaiken, ed. and Y. Trope (eds), *Dual-process Theories in Social Psychology* , pp. 231–254. New York: Guilford.

Fiske, S.T. and Neuberg, S.L (1990) A continuum model of impression formation, from category-based to individuating processes: Influence of information and motivation on attention and interpretation . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 23, 1–74. New York: Academic Press.

Fiske, S.T., Neuberg, S.L., Beattie, A.E. and Milberg, S.J. (1987) Category-based and attribute-based reactions to others: Some informational conditions of stereotyping and individuating processes . *Journal of Experimental Social Psychology* , 23, 399–427.
<http://dx.doi.org/10.1016/0022-1031%2887%2990038-2>

Fiske, S.T. and Pavelchak, M.A. (1986) Category-based versus piecemeal-based affective responses: Developments in schema-triggered affect . In R.M. Sorrentino, ed. and E.T. Higgins (eds), *Handbook of Motivation and Cognition: Foundations of Social Behavior* , pp. 167–203. New York: Guilford Press.

Fiske, S.T. and Ruscher, J.B. (1993) Negative interdependence and prejudice: Whence the affect? In D.M. Mackie, ed. and D.L. Hamilton (eds), *Affect, Cognition, and Stereotyping: Interactive Processes in Group Perception* , pp. 239–268. New York: Academic Press.

Fiske, S.T. and Taylor, S.E. (2008) *Social Cognition: From Brains to Culture* . New York: McGraw-Hill.

Fiske, S.T. and Von Hendy, H.M. (1992) Personality feedback and situational norms can control stereotyping processes . *Journal of Personality and Social Psychology* , 62, 577–596. <http://dx.doi.org/10.1037/0022-3514.62.4.577>

Galinsky, A.D., Gruenfeld, D.H. and Magee, J.C. (2003) From power to action . *Journal of Personality and Social Psychology* , 85, 453–466. <http://dx.doi.org/10.1037/0022-3514.85.3.453>

Glick, P. and Fiske, S.T. (1996) The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism . *Journal of Personality and Social Psychology* , 70, 491–512.
<http://dx.doi.org/10.1037/0022-3514.70.3.491>

Glick, P., Fiske, S.T., Mladinic, A., Saiz, J.L., Abrams, D., Masser, B., et al. (2000) Beyond prejudice as simple antipathy: Hostile and benevolent sexism across cultures . *Journal of Personality and Social Psychology* , 79, 763–775. <http://dx.doi.org/10.1037/0022-3514.79.5.763>

Goodwin, S.A., Fiske, S.T., Rosen, L.D. and Rosenthal, A.M. (2002) The eye of the beholder: Romantic goals and impression biases . *Journal of Experimental Social Psychology* , 38, 232–241. <http://dx.doi.org/10.1006/jesp.2001.1508>

Goodwin, S.A., Gubin, A., Fiske, S.T. and Yzerbyt, V. (2000) Power can bias impression formation: Stereotyping subordinates by default and by design . *Group Processes and Intergroup Relations* , 3, 227–256. <http://dx.doi.org/10.1177/1368430200003003001>

Guinote, A. (2008) Power and affordances: When the situation has more power over powerful than powerless individuals . *Journal of Personality and Social Psychology* , 95, 237–252. <http://dx.doi.org/10.1037/a0012518>

Guinote, A., Brown, M. and Fiske, S.T. (2006) Minority status decreases sense of control and increases interpretive processing . *Social Cognition* , 24, 169–186. <http://dx.doi.org/10.1521/soco.2006.24.2.169>

Guinote, A. and Fiske, S.T. (2003) Being in the out-group territory increases stereotypic perceptions of outgroups: Situational sources of category activation . *Group Processes and Intergroup Relations* , 5, 323–331. <http://dx.doi.org/10.1177/13684302030064001>

Guthrie, E.R. (1935) *The Psychology of Learning* . Oxford: Harper.

Hamilton, D.L and Zanna, M.P. (1974) Context effects in impression formation: Changes in connotative meaning . *Journal of Personality and Social Psychology* , 29, 649–654. <http://dx.doi.org/10.1037/h0036633>

Harris, L.T. and Fiske, S.T. (2006) Dehumanizing the lowest of the low: Neuro-imaging responses to extreme outgroups . *Psychological Science* , 17, 847–853. <http://dx.doi.org/10.1111/j.1467-9280.2006.01793.x>

Hart, A.J., Whalen, P.J., Shin, L.M., McInerney, S.C., Fischer, H. and Rauch, S.L. (2000) Differential response in the human amygdala to racial outgroup vs ingroup face stimuli . *Neuroreport: For Rapid Communication of Neuroscience Research* , 11, 2351–2355. <http://dx.doi.org/10.1097/00001756-200008030-00004>

Hastie, R., ed. , Ostrom, T.M., ed. , Ebbesen, E.B., ed. , WyerJr., R.S., ed. , Hamilton, D.L, ed. and Carlston, D.E. (eds) (1980) Person Memory: The Cognitive Basis of Social Perception . Hillsdale, NJ: Erlbaum.

Hastie, R. and Park, B. (1986) The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line . *Psychological Review* , 93, 258–268. <http://dx.doi.org/10.1037/0033-295X.93.3.258>

Heider, F. (1958) The Psychology of Interpersonal Relations . Hoboken, NJ: Wiley. <http://dx.doi.org/10.1037/10628-000>

Heilman, M.E. and Haynes, M.C. (2008) Subjectivity in the appraisal process: A facilitator of gender bias in work settings . In E. Borgida, ed. and S.T. Fiske (eds), *Psychological Science in the Courtroom: Beyond Common Sense* , pp. 127–156. London: Blackwell.

Kang, J. and Banaji, M. (2006) Fair measures: A behavioral realist revision of ‘affirmative action’ . *California Law Review* , 94, 1063–1118. <http://dx.doi.org/10.2307/20439059>

Katz, I. and Hass, R.G. (1988) Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures . *Journal of Personality and Social Psychology* , 55, 893–905. <http://dx.doi.org/10.1037/0022-3514.55.6.893>

Keltner, D., Gruenfeld, D.H. and Anderson, C. (2003) Power, approach, and inhibition . *Psychological Review* , 110, 265–284. <http://dx.doi.org/10.1037/0033-295X.110.2.265>

Kervyn, N., Fiske, S.T. and Yzerbyt, V. (submitted) Why is the primary dimension of social cognition so hard to predict? Symbolic and realistic threats together predict warmth in the stereotype content model. Unpublished manuscript.

Krieger, L.H. and Fiske, S.T. (2006) Behavioral realism in employment discrimination law: Implicit bias and disparate treatment . *California Law Review* , 94, 997–1062. <http://dx.doi.org/10.2307/20439058>

Kunda, Z. and Thagard, P. (1996) Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory . *Psychological Review* , 103, 284–308. <http://dx.doi.org/10.1037/0033-295X.103.2.284>

Lieberman, M.D., Gaunt, R., Gilbert, D.T. and Trope, Y. (2002) Reflexion and reflection: A social cognitive neuroscience approach to attributional inference . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 34, 199–249. San Diego: Academic Press.

Lin, M.H., Kwan, V.S.Y., Cheung, A. and Fiske, S.T. (2005) Stereotype content model explains prejudice for an envied outgroup: Scale of Anti-Asian American Stereotypes . *Personality and Social Psychology Bulletin* , 31, 34–47. <http://dx.doi.org/10.1177/0146167204271320>

Markus, H. (1977) Self-schemata and processing information about the self . *Journal of Personality and Social Psychology* , 35, 63–78. <http://dx.doi.org/10.1037/0022-3514.35.2.63>

Mather, M., Johnson, M.K. and De Leonardis, D.M. (1999) Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates . *Cognitive Neuropsychology* , 16, 437–458. <http://dx.doi.org/10.1080/026432999380870>

Neuberg, S.L. (1989) The goal of forming accurate impressions during social interactions: Attenuating the impact of negative expectancies . *Journal of Personality and Social Psychology* , 56, 374–386. <http://dx.doi.org/10.1037/0022-3514.56.3.374>

Neuberg, S.L. and Fiske, S.T. (1987) Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes . *Journal of Personality and Social Psychology* , 53, 431–444. <http://dx.doi.org/10.1037/0022-3514.53.3.431>

Neuberg, S.L. and Newsom, J.T. (1993) Personal need for structure: Individual differences in the desire for simpler structure . *Journal of Personality and Social Psychology* , 65, 113–131. <http://dx.doi.org/10.1037/0022-3514.65.1.113>

North, M.S. and Fiske, S.T. (submitted) A new approach to ageism: Succession, identity, and consumption.

Oldmeadow, J. and Fiske, S.T. (2007) Ideology moderates status = competence stereotypes: Roles for belief in a just world and social dominance orientation . *European Journal of Social Psychology* , 37, 1135–1148. <http://dx.doi.org/10.1002/ejsp.428>

Operario, D. and Fiske, S.T. (2001) Effects of trait dominance on powerholders' judgments of subordinates . *Social Cognition* , 19, 161–180. <http://dx.doi.org/10.1521/soco.19.2.161.20704>

Ostrom, T.M. (1977) Between-theory and within-theory conflict in explaining context effects in impression formation . *Journal of Experimental Social Psychology* , 13, 492–503. <http://dx.doi.org/10.1016/0022-1031%2877%2990033-6>

Overbeck, J.R. and Park, B. (2001) When power does not corrupt: Superior individuation processes among powerful perceivers . *Journal of Personality and Social Psychology* , 81, 549–565. <http://dx.doi.org/10.1037/0022-3514.81.4.549>

Pavelchak, M.A. (1989) Piecemeal and category-based evaluation: An idiographic analysis . *Journal of Personality and Social Psychology* , 56, 354–363. <http://dx.doi.org/10.1037/0022-3514.56.3.354>

Phelps, E.A., O'Connor, K.J., Cunningham, W.A., Funayama, E.S., Gatenby, J.C., Gore, J.C. and Banaji, M.R. (2000) Performance on indirect measures of race evaluation predicts amygdala activation . *Journal of Cognitive Neuroscience* , 12, 729–738. <http://dx.doi.org/10.1162/089892900562552>

Rosenberg, S. and Sedlak, A. (1972) Structural representations of implicit personality theory . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 6, 235–297. New York: Academic Press.

Rudman, L.A., Glick, P. and Phelan, J.E. (2008) From the laboratory to the bench: Gender stereotyping research in the courtroom . In E. Borgida, ed. and S.T. Fiske (eds), *Psychological Science in the Courtroom: Beyond Common Sense* , pp. 83–103. London: Blackwell.

Ruscher, J.B. and Fiske, S.T. (1990) Interpersonal competition can cause individuating processes . *Journal of Personality and Social Psychology* , 58, 832–843. <http://dx.doi.org/10.1037/0022-3514.58.5.832>

Ruscher, J.B., Fiske, S.T., Miki, H. and Van Manen, S. (1991) Individuating processes in competition: Interpersonal versus intergroup . *Personality and Social Psychology Bulletin* , 77, 595–605. <http://dx.doi.org/10.1177/0146167291176001>

Russell, A.M. and Fiske, S.T. (2008) It's all relative: Social position and interpersonal perception . *European Journal of Social Psychology* , 38, 1193–1201. <http://dx.doi.org/10.1002/ejsp.539>

Snyder, M. and Swann, W.B. (1978) Behavioral confirmation in social interaction: From social perception to social reality . *Journal of Experimental Social Psychology* , 14, 148–162. <http://dx.doi.org/10.1016/0022-1031%2878%2990021-5>

Srull, T.K. (1981) Person memory: Some tests of associative storage and retrieval models . *Journal of Experimental Psychology: Human Learning and Memory* , 7, 440–462. <http://dx.doi.org/10.1037/0278-7393.7.6.440>

Stangor, C. and McMillan, D. (1992) Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures . *Psychological Bulletin* , 111, 42–61. <http://dx.doi.org/10.1037/0033-2909.111.1.42>

Stevens, L.E. and Fiske, S.T. (1995) Motivation and cognition in social life: A social survival perspective . *Social Cognition* , 13, 189–214. <http://dx.doi.org/10.1521/soco.1995.13.3.189>

Stevens, L.E. and Fiske, S.T. (2000) Motivated impressions of a powerholder: Accuracy under task dependency and misperception under evaluative

dependency . *Personality and Social Psychology Bulletin* , 26, 907–922. <http://dx.doi.org/10.1177/01461672002610002>

Talaska, C.A., Fiske, S.T. and Chaiken, S. (2008) Legitimating racial discrimination: A meta-analysis of the racial attitude-behavior literature shows that emotions, not beliefs, best predict discrimination . *Social Justice Research: Social Power in Action* , 21, 263–296. <http://dx.doi.org/10.1007/s11211-008-0071-2>

Taylor, S.E. and Fiske, S.T. (1978) Salience, attention, and attribution: Top-of-the-head phenomena . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 11, 249–288. New York: Academic Press.

Taylor, S.E., Fiske, S.T., Etcoff, N.L. and Ruderman, A.J. (1978) Categorical and contextual bases of person memory and stereotyping . *Journal of Personality and Social Psychology* , 36, 778–793. <http://dx.doi.org/10.1037/0022-3514.36.7.778>

Wheeler, M.E. and Fiske, S.T. (2005) Controlling racial prejudice and stereotyping: Social cognitive goals affect amygdala and stereotype activation . *Psychological Science* , 16, 56–63. <http://dx.doi.org/10.1111/j.1467-9280.2006.01793.x>

Willis, J. and Todorov, A. (2006) First impressions: Making up your mind after a 100-ms exposure to a face . *Psychological Science* , 17, 592–598. <http://dx.doi.org/10.1111/j.1467-9280.2006.01750.x>

Wojciszke, B. (2005) Morality and competence in person- and self-perception . *European Review of Social Psychology* , 16, 155–188. <http://dx.doi.org/10.1080/10463280500229619>

<http://dx.doi.org/10.4135/9781446249215.n14>

Handbook of Theories of Social Psychology: Volume 1

Feelings-as-Information Theory

Contributors: Norbert Schwarz

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Feelings-as-Information Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n15>

Print pages: 289-309

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n15>

[p. 289 ↓]

Chapter 14: Feelings-as-Information Theory

Norbert Schwarz, ed.

Abstract

Feelings-as-information theory conceptualizes the role of subjective experiences – including moods, emotions, metacognitive experiences, and bodily sensations – in judgment. It assumes that people attend to their feelings as a source of information, with different feelings providing different types of information. Whereas feelings elicited by the target of judgment provide valid information, feelings that are due to an unrelated influence can lead us astray. The use of feelings as a source of information follows the same principles as the use of any other information. Most important, people do not rely on their feelings when they (correctly or incorrectly) attribute them to another source, thus undermining their informational value for the task at hand. What people conclude from a given feeling depends on the epistemic question on which they bring it to bear; hence, inferences from feelings are context-sensitive and malleable. In addition to serving as a basis of judgment, feelings inform us about the nature of our current situation and our thought processes are tuned to meet situational requirements. The chapter reviews the development of the theory, its core propositions and representative findings

Introduction

Human thinking is accompanied by a variety of subjective experiences, including moods and emotions, metacognitive feelings (like ease of recall or fluency of perception),

and bodily sensations. Feelings-as-information theory provides a general framework for conceptualizing the role of these experiences in human judgment. It was initially developed to account for the influence of happy and sad moods on evaluative judgment. However, the theoretical principles of the initial mood-as-information work (Schwarz and Clore, 1983) could be fruitfully applied to other types of feelings and developed into a more comprehensive conceptualization of the interplay of feeling and thinking. This chapter summarizes what has been learned.

The first section provides a short personal account of the theory's development and places its assumptions in their historical context (for more detailed discussions of other theoretical approaches, see Clore et al., 1994; Schwarz and Clore, 2007). The second section [p. 290 ↓] presents the theory's postulates and the third section reviews representative findings.

A Look Back: The Development of Feelings-as-Information Theory

As many readers know from personal experience, our lives look better on some days than on others, even though nothing of any obvious importance has changed. In my case, the upbeat or gloomy mood induced by a sunny or rainy day is sufficient to do the trick. Trying to understand this experience as a graduate student at the University of Mannheim, Germany, in the late 1970s, I turned to what social and cognitive psychologists had learned from the experimental mood research available at the time. One account, advanced by Isen et al. (1978) and Bower (1981), held that moods increase the accessibility of mood-congruent information in memory. From this perspective, positive (negative) aspects of life are more likely to come to mind when we are in a happy (sad) mood, resulting in mood-congruent judgments. This approach was consistent with social psychology's new adoption of the information processing paradigm and its emphasis on storage and retrieval processes. However, it didn't seem "quite right" introspectively: on good days, things just "felt" better and this did not seem to involve selective recall of past events of mood-congruent valence. Phenomenological analyses in the introspective tradition of German "armchair psychology" (e.g., Bollnow, 1956), which treated moods as an integrative reflection of one's current situation,

seemed closer to the mark – alas, such introspections are to be taken with a grain of salt (Nisbett and Wilson, 1977). A competing perspective, Zajonc's (1980) "affective primacy" hypothesis, had the advantage of avoiding reliance on mood-congruent retrieval processes but lacked a process model specific enough to meet the developing criteria of social cognition research.

A conversation with Bob Wyer offered a different approach. Wyer and Carlston (1979) proposed that affect can serve informational functions, "for example, one's liking for a person may be based partly on the feelings of pleasantness when the person is around" (1979: 192). In addition, they conjectured that affective states may direct our attention to information that is suitable to explain one's feelings. While their conjectures were compatible with phenomenological approaches, their conceptualization emphasized the role of cognitive representations of experience at the expense of actual current experience itself, consistent with the information processing paradigm. Research into the influence of arousal (from Schachter and Singer's emotion research [1962] to Zillman's arousal-transfer model [1978] and Zanna and Cooper's dissonance studies [1976]) suggested, however, that the online experience itself may play a crucial role. More importantly, this literature also suggested that misattribution manipulations would be suitable experimental tools to address the role of current experience in human judgment.

A post-doctoral year with Wyer and Clore at the University of Illinois provided the opportunity to pursue these issues. Clore and Byrne (1974) had proposed a reinforcement-affect model to account for affective influences on interpersonal attraction. Going beyond the learning theories of the time, their model assumed that rewards exert their influence through the positive affect they elicit. Supporting this notion, laboratory and field experiments showed that associating others with positive feelings is sufficient to increase interpersonal attraction, even when the feelings are incidental and due to an unrelated source (e.g., Griffitt and Veitch, 1971). By the late 1970s Clore began to wonder why we "don't all end up falling in love with the paymaster," as he put it, if the mere co-occurrence of people with reward is sufficient to induce attraction. Does incidental affect only influence our judgments when we are not aware of its source, as Zillman's (1978) arousal studies suggested?

[p. 291 ↓]

Initial Evidence

These converging interests resulted in a conceptually straightforward study (Schwarz and Clore, 1983: Experiment 1). We asked participants to vividly recall and describe a happy or sad event to induce a corresponding mood and crossed these mood inductions with a misattribution manipulation that took advantage of a somewhat bizarre little room, previously used for auditory research with monkeys (for the inside story see Schwarz and Clore, 2003). This allowed us to suggest to some participants that the room may induce elated feelings and to others it may induce depressed feelings. Judgments of life-satisfaction served as the dependent variable. Our procedure deliberately stacked the deck in favor of content-driven models: by inducing moods through the recall of a happy or sad event, mood-congruent recall would be facilitated both by the content of the recall task and the induced mood. The predictions were straightforward. If mood effects on judgment were a function of mood-congruent recall (Bower, 1981), participants should report higher life-satisfaction when in a happy rather than sad mood, independent of what we told them about the room. If the experience itself served informative functions, on the other hand, its impact should depend on the feeling's perceived diagnosticity. That is, mood effects on reported life-satisfaction should be attenuated when the mood is attributed to the influence of the room and hence considered uninformative for evaluating one's life in general (a discounting effect in Kelley's [1972] terms); however, it should be enhanced when one experiences the mood despite an allegedly opposing influence of the room (an augmentation effect).

The results, shown in [Table 14.1](#), were consistent with these predictions, although only under sad mood conditions. The latter observation was compatible with Wyer and Carlston's (1979) suggestion that sad moods may require more explanation than happy ones, which would render them more susceptible to attributional manipulations. More important, the obtained augmentation and discounting effects made it unlikely that the influence of moods was driven by mood-congruent recall. After all, we had induced moods by having participants recall happy or sad events, thus adding semantic priming to the assumed affective activation of valenced material (Bower, 1981). Nevertheless, the accessible semantic content had little impact when participants discounted the accompanying negative feelings, assigning a crucial role to the subjective experience

itself. Finally, the obtained attributional effects highlighted that the path from feelings to judgment was inferential, in contrast to Zajonc's (1980) assertion that "preferences need no inferences."

Table 14.1 Life-satisfaction as a function of mood and attribution

Induced mood	Expectation about room		
	Tense	None	Elated
Positive	9.6a	8.6a	9.7a
Negative	8.6a	5.7b	4.4b
Control	-	8.9a	-

Note: Shown are mean reports of life-satisfaction (11 = very satisfied). Means not sharing a common subscript differ at $p < 0.05$.

Source: Adopted from Schwarz and Clore (1983).

A subsequent, more naturalistic study took advantage of sunny and rainy weather as a mood manipulation (Schwarz and Clore, 1983: Experiment 2). As daily experience [p. 292 ↓] suggests, participants reported higher life-satisfaction (and a more positive mood) when they were called on sunny than on rainy days. However, the negative influence of bad weather was eliminated when the interviewer, who pretended to call from out of town, first inquired about the weather at respondents' place of residence. This discounting effect was not obtained under sunny weather conditions, again suggesting that sad moods are more likely to be explained than happy moods. In combination, these studies provided first evidence for several assumptions that became core themes in the development of feelings-as-information theory.

Core Themes

First, our findings showed that people attend to their momentary feelings as a source of information in forming judgments, essentially asking themselves, "How do I feel about this?" Later research extended this "informative function" (Wyer and Carlston, 1979) of affective states to other feelings, including nonaffective feelings, like the metacognitive experience of ease of recall (Schwarz et al., 1991b), and bodily sensations (Stepper and Strack, 1993).

Second, the observed discounting and augmentation effects highlighted that people use their feelings like any other source of information. They do not rely on them when they become aware that their feelings may be due to an unrelated source, thus undermining their informational value for the judgment at hand. Conversely, they consider their feelings particularly informative when they experience them despite opposing forces. Later research, much of it conducted by Michel Pham and his colleagues, identified additional variables that influence how much weight we give to our feelings (see Pham, 2004).

Third, our initial studies documented more positive judgments under happy than sad moods. While this is true for the bulk of mood research (Schwarz and Clore, 2007), Leonard Martin and colleagues (e.g., Martin et al., 1997) demonstrated that positive feelings can result in negative evaluations. For example, when we feel happy while reading a sad story, we may conclude that it is not a “good sad story” after all, or else it would make us feel sad. Such findings illustrate that the influence of feelings depends on the specific question on which the feeling is brought to bear. This theme proved particularly important in later research on metacognitive experiences (Schwarz, 2004).

Finally, the observation that misattribution effects only emerged under sad moods ([Table 14.1](#)) proved more puzzling. Because most people feel mildly positive most of the time (Matlin and Stang, 1979), we initially suggested that sad moods are deviations from one's usual state and hence more likely to require explanation. This, in turn, would direct attention to possible sources of one's mood (Wyer and Carlston, 1979), rendering sad moods more susceptible to (mis)attribution manipulations. If so, being in an unexplained sad mood should interfere with other cognitive tasks, due to the competing demands of explaining one's mood. Testing this prediction, Bless et al. (1990) exposed participants in happy or sad moods to strong or weak persuasive arguments and assumed that sad moods would reduce systematic message elaboration. To our surprise, we found the opposite: sad participants engaged in message elaboration, whereas happy participants did not, by now a familiar and frequently replicated finding (for a review see Schwarz et al., 1991a). Similarly, Sinclair (1988) reported strong evidence that being in a sad mood reduced halo effects in impression formation. Clearly, sad moods did not pose an explanation problem that interfered with other processing demands; to the contrary, sad moods increased, and happy moods decreased, systematic processing in these studies.

To account for these findings, we (Schwarz, 1987, 1990; Schwarz and Bless, 1991) suggested that the informative function of moods is more general than assumed in the initial theorizing, which had focused on evaluative judgment. [p. 293 ↓] In daily life, we usually feel bad when we encounter a threat of negative or a lack of positive outcomes, and feel good when we obtain positive outcomes and are not threatened by negative ones. Hence, our moods reflect the state of our environment (Bollnow, 1956). If so, bad moods may signal a problematic situation, whereas good moods may signal a benign situation. Given the situated nature of human cognition, we may expect that our thought processes are tuned to meet the processing requirements apparently posed by the situation, resulting in systematically different processing strategies under happy and sad moods. Sad moods may foster a systematic processing style that is characterized by bottom-up processing, attention to the details at hand, and limited playfulness and creativity. Happy moods, on the other hand, may foster a top-down processing style that relies more on general knowledge structures and is accompanied by less focused attention and higher playfulness and creativity. By and large these assumptions proved compatible with the accumulating evidence (for reviews see Clore et al., 1994; Schwarz, 2002; Schwarz and Clore, 2007). Moreover, later research showed that any information that signals a benign or problematic situation – from bodily sensations (e.g., Friedman and Förster, 2000) to metacognitive experiences (Song and Schwarz, 2008) or the smiling or frowning face of a communicator (Ottati et al., 1997) – can elicit the corresponding processing style. From this perspective, the misattribution effects observed by Schwarz and Clore (1983) were limited to sad moods because sad moods facilitate the analytic reasoning needed for attributional analyses, whereas happy moods make such reasoning less likely.

In the following sections, I revisit these themes by reviewing the postulates of feelings-as-information theory and illustrative experimental evidence. If social psychologists followed the naming traditions of software engineers, this would arguably be FAIT.3. Its treatment of the use of feelings as a source of information in judgment differs from FAIT.1 (Schwarz and Clore, 1983) by emphasizing that the specific impact of a feeling depends on the epistemic question on which it is brought to bear. Its treatment of the influence of feelings on processing style differs from FAIT.2 (Schwarz, 1990; Schwarz and Bless, 1991) by de-emphasizing the role of processing ability (consistent with Bless and Schwarz, 1999) and by extending the range of variables that influence processing

style beyond the role of feelings (consistent with Schwarz, 2002). Moreover, the theory's treatment of feelings goes far beyond our initial emphasis on moods and emotions and includes nonaffective experiences, consistent with the work conducted in the quarter century since the initial studies.

Feelings-as-Information Theory: Postulates

The core postulates are summarized in [Box 14.1](#); they bear on the perceived informational value of feelings, their use as a basis of judgment, and their influence on the spontaneous adoption of different processing styles.

What Feelings Convey

The theory postulates that people attend to their feelings as a source of information, with different types of feelings providing different types of information. This assumption has a long tradition in emotion research. As Frijda (1988: 354) put it, "emotions exist for the sake of signaling states of the world that have to be responded to, or that no longer need response and action." What exactly a given emotion signals can be derived from its underlying appraisal pattern (Ellsworth and Scherer, 2003; Ortony et al., 1988). Anger, for example, is a response to a loss or lack of reward that is attributed to the causal action of another agent; when no agent attribution is [[p. 294 ↓](#)] made, a loss gives rise to sadness. Accordingly, anger and sadness not only inform us about a loss, but also about its cause and elicit downstream responsibility judgments that reflect this information (e.g., Keltner et al., 1993a). Because emotions arise from ongoing, implicit appraisals of situations with respect to their implications for one's goals, they have an identifiable referent (what the emotion is "about"), a sharp rise time, and limited duration. These characteristics distinguish emotions from *moods*, which lack a clear referent, may come about gradually, may last for an extended time, and are often of low intensity (Bollnow, 1956; Morris, 1989). Hence, moods are more diffuse than emotions and primarily convey generic valence information that lacks a clear referent. These differences are apparent when we say that we are angry "about" something, but "in" a bad mood.

Box 14.1 Postulates

- 1. People attend to their feelings as a source of information.
 - a. Different types of feelings provide different types of information.
- 2. The impact of a given feeling depends on its perceived informational value for the task at hand.
 - a. People usually experience their feelings as being “about” whatever is in the focus of attention; this fosters the perception that incidental feelings are relevant.
 - b. When a feeling is attributed to an incidental source, its informational value is discounted; conversely, when it is experienced despite perceived opposing forces, its informational value is augmented.
 - c. Changes in one's feelings are more informative than stable states.
- 3. When feelings are used as information, their use follows the same principles as the use of any other type of information.
 - a. The impact of feelings increases with their perceived relevance to the task at hand and decreases with the accessibility and consideration of alternative diagnostic inputs, which is a function of processing motivation and capacity.
 - b. What people conclude from a given feeling depends on (i) the epistemic question on which they bring it to bear and (ii) the lay theory applied.
- 4. Like other information, feelings can
 - a. serve as a basis of judgment
 - b. influence the choice of processing strategies; feelings that signal a “problematic” situation foster an analytic, bottom-up processing style, whereas feelings that signal a “benign” situation foster a more global, top-down processing style.

Cognitive feelings like surprise, boredom, or feelings of familiarity provide information about the state of one's knowledge (Ortony et al., 1988). Of particular interest to social psychologists is the metacognitive experience of ease or difficulty, which can pertain to recall and thought generation (*accessibility experiences*; Schwarz, 1998) or to the processing of new, external information (*processing fluency*, Winkielman et

al., 2003). Numerous variables can influence these experiences, from the amount of information a person tries to recall to the presentation format in which new information is presented (e.g., print fonts, figure-ground contrast) and the semantic context in which it is embedded. Because cognitive operations can be easy or difficult for many different reasons, the specific inferences people draw from these experiences depend on which of many lay theories of mental processes they [p. 295 ↓] bring to bear on the task (Schwarz, 2004). In addition, easy processing is experienced as pleasant (as captured by psychophysiological measures; Winkielman and Cacioppo, 2001) and this affective response can itself serve as a basis of judgment (Winkielman et al., 2003).

Finally, *bodily experiences* include feelings like hunger, pain, and physiological arousal, which inform us about physical states of the organism. Other bodily experiences provide information that parallels the implications of affective and cognitive feelings. For example, furrowing one's brow (contraction of the zygomaticus) conveys a feeling of effort and affects judgment in ways that parallel the metacognitive experience of difficulty (e.g., Sanna et al., 2002; Stepper and Strack, 1993). Similarly, proprioceptive feedback from facial expressions (e.g., Strack et al., 1988) and arm flexion and extension (e.g., Friedman and Förster, 2000) influence judgment and processing style in ways that parallel affective influences.

Perceived Informational Value

The theory further postulates that the impact of a given feeling depends on its perceived informational value for the task at hand. When a feeling is elicited by the object of judgment ("integral" in Bodenhausen's [1993] terminology), it provides valid information about the person's own response to the target. For example, seeing Susan may elicit positive feelings in Tom and he may be well advised to consider these feelings in (some) judgments of Susan. When the feeling is due to some other source ('incidental'), however, it provides (potentially) misleading information; for example, Tom's good feelings may be due to the weather rather than Susan. Unfortunately, people are more sensitive to their feelings than to where their feelings come from. They commonly assume that any feelings they have, and any thoughts that come to mind, are "about" whatever is in the focus of their attention (Higgins, 1998) – or why else would they have them now, in this context? Hence, they are likely to perceive incidental feelings as being

“about” the target of judgment, unless their attention is drawn to a plausible incidental source.

Whenever people (correctly or incorrectly) attribute their feelings to an incidental source, the perceived informational value of their feelings for the judgment at hand is undermined. Conversely, when they perceive that they have these feelings despite opposing forces, their feelings' perceived informational value is augmented. The sad mood conditions of [Table 14.1](#) illustrate these discounting and augmentation effects.

When the informational value of their feelings is called into question, people turn to other sources of information to arrive at a judgment. As seen above, participants in the Schwarz and Clore (1983) study who discounted their sad mood arrived at life-satisfaction judgments that did not differ from participants in the control condition, who were not exposed to a mood manipulation. Presumably, both groups could draw on extensive other information about their own lives, resulting in similar judgments. Had such alternative inputs not been available, they might have resorted to an inferential correction strategy akin to, “I feel bad about my life, but this may be due to the room – so I should adjust my judgment upward.” Such theory-driven correction strategies usually result in overcorrection; that is, a bias in the opposite direction (Strack and Hannover, 1996; Wilson and Brekke, 1994). Accordingly, discounting one's feelings as a source of information can either eliminate their influence (when alternative sources of information are accessible) or elicit a bias in the opposite direction (due to overcorrection in the absence of alternative inputs).

The theory further assumes that changes in one's feelings are more informative than stable states. This assumption is consistent with numerous studies in sensory perception and the covariation principle of attribution research. However, it has rarely been explicitly [p. 296 ↓] tested in feelings-as-information experiments (for exceptions see Hansen et al., 2008; Shen et al., 2010). By relying on the experimental induction of feelings, experiments always involve a change from baseline as part of the methodological routine, which contributes to the feeling's perceived informational value.

Some Misunderstandings

Some common misunderstandings of these assumptions deserve attention. To disentangle the contributions of the perceiver's feelings from other information about the target, experimental tests of the feelings-as-information hypothesis rely on the induction of incidental feelings. This gave rise to the erroneous conclusion that the use of feelings as a source of information is limited to incidental feelings, which ledForgas (2001: 104) to assert that "affect can only serve as a heuristic cue due to mistaken inferences," making reliance on one's feelings "an ineffective and dysfunctional strategy." This assertion confuses the operational and theoretical level. While reliance on incidental feelings can indeed be dysfunctional, integral feelings provide valid information. Attending to this information is highly adaptive, as a large body of research on emotional intelligence and the role of feelings in decision making indicates (see Barrett and Salovey, 2002; Damasio, 1994).

Falling prey to the same confusion, Slovic and colleagues (see Slovic et al., 2002) proposed an "affect heuristic" to account for the influence of integral feelings, which they considered distinct from the influence of incidental feelings. Unfortunately, integral feelings are inherently confounded with the positive or negative target attributes that elicit them, making it impossible to determine if observed differences are driven by experiential information in the form of integral feelings (as Slovic and colleagues assume) or by declarative information in the form of different target attributes. From the perspective of feelings-as-information theory, the use of integral and incidental feelings as a source of information reflects the same basic mechanism – and any influence of target attributes that is not mediated by the feelings they elicit is better described in terms of declarative rather than "affective" information.

Finally, some observers (e.g., Forgas, 2001) suggested that feelings-as-information effects require a conscious attribution of the feeling to the target. This is not the case. Whereas discounting and augmentation effects require some level of conscious attribution, the mere use of one's feelings as a source of information does not. As noted, people usually consider their thoughts and feelings to be "about" whatever is in the focus of their attention, rendering reliance on them the automatic default option. Accordingly, the impact of feelings increases when contextual influences, like time

pressure (Siemer and Reisenzein, 1998), limit the opportunity to engage in attributional analyses, in contrast to what a conscious attribution requirement would predict.

From Feelings to Judgments

The theory further postulates that whenever feelings are used as a source of information, their use follows the same rules as the use of any other information. First, feelings are only used as a source of information when their informational value is not called into question (e.g., Schwarz and Clore, 1983). Second, the impact of feelings increases with their perceived relevance to the judgment at hand. For example, moods exert a stronger influence when people make decisions for themselves rather than others, whose affective response may differ from their own (Raghunathan and Pham, 1999) or when they evaluate the hedonic pleasure that can be derived from an activity rather than the activity's instrumental value for academic achievement (Pham, 1998). Third, the impact of feelings decreases the more other relevant [p. 297 ↓] inputs are accessible. For example, people are less likely to rely on their feelings when they have high expertise in the domain of judgment (e.g., Ottati and Isbell, 1996; Sedikides, 1995), which presumably facilitates the assessment of the relevance of one's feelings and renders other information easily accessible. Fourth, as is the case for any other highly accessible piece of information, the impact of feelings is more pronounced under conditions of low processing capacity (e.g., Greifeneder and Bless, 2007; Siemer and Reisenzein, 1998) or motivation (e.g., Rothman and Schwarz, 1998). These conditions limit assessments of the diagnosticity and relevance of one's feelings and the search for possible alternative inputs. As these examples illustrate, the variables that govern the use and impact of experiential information as a basis of judgment parallel the variables that govern the use and impact of declarative information, consistent with the basic feelings-as-information logic.

Finally, feelings share with other information that their specific implications depend on the question asked. The observation that Bob has published a highly acclaimed book every year since his PhD can be brought to bear on many judgments of Bob, from his intelligence and ambitiousness to his professional standing and his commitment to spending time with his kids. The same holds for feelings. What people conclude from a given feeling depends on the epistemic question on which they bring it to bear. For

example, Martin et al. (1993) asked happy and sad participants to list birds. When asked whether they are satisfied with what they accomplished, happy participants inferred that they are satisfied and terminated the task, whereas sad participants inferred that they are not yet satisfied and continued. This pattern reversed when participants were asked whether they enjoy what they are doing. In this case, happy participants inferred enjoyment and continued with the task, whereas sad participants inferred a lack of enjoyment and terminated the task. In both cases, their judgments were consistent with the valence information provided by their mood, yet this valence information had diverging behavioral implications, depending on the specific question on which it was brought to bear.

Importantly, some feelings require more interpretation, and allow for a wider range of inferences, than others. As already noted, moods provide broadly applicable valence information, whereas specific emotions inform us that a specific appraisal pattern has been met, which constrains the range of plausible inferences. At the other extreme, meta-cognitive experiences primarily inform us that our cognitive operations are easy or difficult – and they may be so for many reasons. For example, we may find it difficult to recall information because the event happened a long time ago, because we never found it important and hence didn't pay attention, because we lack expertise in the domain, and so on. Which inferences we draw from difficulty of recall will therefore depend on which of these naïve theories of mental processes we bring to bear.

Applicable theories are usually brought to mind by the judgment task (Schwarz, 2004) and the same metacognitive experience can result in differential judgments of expertise, importance or temporal distance, depending on the specific question asked.

Cognitive Tuning: Feelings and Processing Style

In addition to providing information that can serve as a basis of judgment, feelings influence *how* people process information; that is, their processing style. A number of different explanations have been offered for this observation, usually highlighting the role of one specific type of feeling (for reviews, see Schwarz and Clore, 2007, and the contributions in Martin and Clore, 2001). Feelings-as-information theory provides

a unified conceptualization of these influences in the context of a situated cognition framework (Smith and Semin, 2004). It assumes that human cognition stands in the service of [p. 298 ↓] action (James, 1890) and that our cognitive processes are responsive to the environment in which we pursue our goals. This responsiveness ranges from the higher accessibility of knowledge relevant to the current situation (e.g., Yeh and Barsalou, 2006) to the choice of processing strategies that meet situational requirements (e.g., Wegner and Vallacher, 1986). When things go smoothly and we face no hurdles in the pursuit of our goals, we are likely to rely on our pre-existing knowledge structures and routines, which served us well in the past. Moreover, we may be willing to take some risk in exploring novel solutions. Once things go wrong, we abandon reliance on our usual routines and focus on the specifics at hand to determine what went wrong and what can be done about it.

Feelings play a crucial role in this tuning process by providing a fast and parsimonious indicator of whether our current situation is “benign” or “problematic.” The influence of feelings on processing style is eliminated when the informational value of the feeling is called into question (e.g., Sinclair et al., 1994) and can be overridden by the individual's goals or explicit task demands (e.g., Bless et al., 1990).

Representative Findings

Next, I review representative findings pertaining to the influence of moods, emotions, and metacognitive experiences on judgment and processing style and highlight some real-world implications (for more extensive reviews of findings see Clore et al., 1994; Schwarz and Clore, 2007).

Feelings as a Basis of Judgment

Moods

As discussed, moods convey valence information that usually results in more positive judgments when people are in a happy rather than sad mood, with neutral moods falling

in between. This influence is not observed when the informational value of the mood is called into question through (mis)attribution manipulations (Schwarz and Clore, 1983; for conceptual replications see Gorn et al., 1993; Savitsky et al., 1998; Siemer and Reisenzeiri, 1998, among others). Even when one's mood is considered informative, its impact depends on its perceived relevance to the judgment at hand (e.g., Pham, 1998) and the accessibility of competing inputs (e.g., Sedikides, 1995), as discussed above. Importantly, mood effects are not limited to inconsequential judgments. Instead, moods have been found to influence highly consequential decisions, from medical school admissions (Redlmeier and Baxter, 2009) to stock market investments. For example, Hirshleifer and Shumway (2003) observed a reliable influence of the weather on stock market returns in 26 countries: the market is more likely to go up when the sun shines in the city that hosts the country's major stock exchange. Presumably, the upbeat mood associated with sunny weather makes investors more optimistic about the future of the economy, paralleling observations in experiments.

Whereas the bulk of the research shows more positive (negative) judgments under happy (sad) mood, moods can also result in mood-incongruent judgments under specific conditions. First, mood incongruent judgments can result from the logic of discounting effects themselves (e.g., Ottati and Isbell, 1996). Suppose, for example, that you are evaluating a job candidate and are aware that you are in a miserable mood due to an earlier event. To which extent are your bad feelings an integral part of your reaction to the candidate and to which extent are they due to the earlier event? If you fully discount your bad feelings, you may arrive at an unduly positive evaluation of the candidate. Second, mood-inducing events can elicit contrast effects in the evaluation of closely related targets by serving as extreme standards of comparison. For example, Schwarz et al. (1987) conducted an experiment in a very pleasant or unpleasant room. Replicating earlier findings, their [p. 299 ↓] student participants reported higher life-satisfaction when the pleasant room induced a positive mood than when the unpleasant room induced a negative mood. When asked about their housing-satisfaction, however, this pattern reversed, presumably because even modest dorm rooms seemed luxurious compared to the salient standard introduced by the unpleasant room. Finally, the target of judgment can carry affective expectations to which one's current feelings are compared; for example, Martin and colleagues (1997) observed that happy participants rated a sad story less favorably than sad participants. Presumably, their happy feelings

implied that the sad story failed to achieve its goal of making them feel sad, leading them to conclude that it was a poor sad story.

Emotions

As observed for moods, the impact of specific emotions is eliminated when they are attributed to an incidental source. For example, Schwarz et al. (1985) found that a fear-arousing communication did not affect participants' attitudes when they attributed their feelings to allegedly arousing side-effects of a pill; conversely, expecting the pill to have tranquilizing effects enhanced the message's impact. However, the informational value of specific emotions differs from the informational value of global moods in ways that can be traced to the role of appraisals.

Recall that emotions reflect the person's appraisal of a specific event (Ellsworth and Scherer, 2003; Ortony et al., 1988), which is in the focus of the person's attention. This makes emotions less likely to be misread as bearing on unrelated targets than is the case for diffuse moods. Indeed, merely labeling one's current feelings with specific emotion terms is sufficient to elicit an event attribution and has been found to be as efficient in eliminating effects on unrelated judgments as a standard misattribution manipulation (Keltner et al., 1993b). Note that this observation has important methodological implications: using detailed emotion terms as manipulation checks invites causal attributions to determine the specific emotion, which can eliminate the expected effect.

Moreover, experiencing an emotion implies that a specific set of appraisal criteria has been met. Anger, for example, informs us that somebody did us wrong and hence provides more specific information than a diffuse negative mood. Accordingly, the influence of emotions can be predicted on the basis of the underlying appraisals (e.g., Lerner and Keltner, 2000). For example, Lerner and colleagues (2003) observed in a national survey during the immediate aftermath of the terrorist attacks of September 11, 2001, that inducing participants to focus on the experienced fear increased risk estimates and plans for precautionary behavior, whereas focusing on the experienced anger did the reverse.

Metacognitive Experiences

Compared with the appraisal information conveyed by emotions, the information conveyed by metacognitive experiences of ease or difficulty is relatively diffuse. All the experience, by itself, conveys is that “this” is easy or difficult – and most cognitive operations can be so for many different reasons. Hence, the same experience can give rise to different inferences, depending on which of many lay theories of mental processes comes to mind (Schwarz, 2004). As observed for moods and emotions, the influence of metacognitive experiences is eliminated when they are attributed to an incidental source (e.g., Schwarz et al., 1991b).

Accessibility experiences As an example, consider the ease or difficulty with which information can be brought to mind. According to most models of judgment, an object should be evaluated more favorably when we recall many rather than few positive attributes; similarly, an event should seem more likely when we generate many rather than few reasons for its occurrence. Empirically, the opposite is often the case. For example, people consider themselves less assertive after recalling many rather than [p. 300 ↓] few examples of their own assertive behavior (Schwarz et al., 1991b); hold an attitude with less confidence after listing many rather than few supporting arguments (Haddock et al., 1999); and consider an event less likely after listing many rather than few reasons for its occurrence (Sanna and Schwarz, 2004). Throughout, their inferences are consistent with the implications of accessible thought content when thought generation is easy (few), but opposite to these implications when thought generation is difficult (many). This pattern reflects a lay theory of mental processes that is at the heart of Tversky and Kahneman's (1973) availability heuristic: “The more exemplars exist, the easier it is to bring some to mind.” Hence, the difficulty of generating many reasons or examples suggests that there aren't many, giving rise to the above conclusions. When participants attribute the experienced difficulty to an incidental influence, like music played in the background, its informational value is undermined and they turn to accessible thought content as an alternative input. In this case, the otherwise observed pattern reverses and they infer, for example, that they are more assertive, the more examples of assertive behaviors they recall (Schwarz et al., 1991). Similarly, yoked participants, who merely read the thoughts generated

by someone else and are hence deprived of the generation experience, are more influenced when their partner lists many rather than few arguments, in contrast to the person who lists them (Wänke et al., 1996). These observations highlight that the thought content by itself is compelling once it is not qualified by a subjective difficulty experience.

Other lay theories hold, for example, that details of recent events are easier to recall than details of distant events, and details of important events easier than details of unimportant ones. Which of these (or many other) theories comes to mind depends on the question posed. Schwarz and Xu (2010) had participants recall details of the Oklahoma City bombing. When first asked to date the event, participants inferred that it was more recent after recalling two rather than ten details; when first asked how important they found the event at the time, they inferred higher importance after recalling two rather than ten details. Thus, the same accessibility experience informed judgments of temporal distance or of importance, depending on the question posed. More important, application of a given theory entails an attribution of the experience to a specific cause (here, recency or importance), which changes the implications of the experience for other judgments (Schwarz, 2004). Accordingly, participants who initially attributed the difficulty of recalling many details to the event's temporal distance subsequently reported that the event was very important to them – after all, they could still recall details even though the event had apparently happened long ago, so it must have been quite important. Conversely, participants who initially attributed difficulty of recall to low personal importance subsequently dated the event as closer in time – after all, they could still recall details despite the event's low personal importance, so it must have been quite recent. Such findings (for a review see Schwarz, 2010) show that inferences from metacognitive experiences are highly malleable, presumably because people are aware that cognitive operations can be easy or difficult for many different reasons, each of which provides a different inference rule.

Paralleling the findings for other feelings, people are more likely to rely on their accessibility experiences under conditions that commonly foster heuristic processing, but turn to accessible content under conditions that commonly foster systematic processing. The latter conditions include high personal relevance (e.g., Rothman and Schwarz, 1998), high need for cognition (e.g., Greifeneder et al., 2010) and being in a sad rather than happy mood (e.g., Ruder and Bless, 2003).

Processing fluency Just like information can be easy or difficult to bring to mind, new information we encounter can be easy or difficult to process. Numerous variables can [p. 301 ↓] influence ease of processing, ranging from presentation characteristics (like figure-ground contrast, print font, or rhyming versus nonrhyming form) to the semantic relatedness of the material and the frequency and recency of previous exposure. Because these variables result in similar phenomenological experiences, the meaning of the experience is open to interpretation. Which interpretation people choose, and which inferences they draw, again depends on the naïve theory they bring to bear (Schwarz, 2004, 2010).

One naïve theory that is of particular importance to social psychological phenomena is the (usually correct) assumption that familiar material is easier to process than unfamiliar material. Hence, fluently processed material seems more familiar than disfluently processed material, even when the fluency experience is solely due to incidental variables, like the print font or color contrast in which the material is presented. As observed for other feelings, drawing people's attention to these incidental sources of fluency undermines the informational value of the experience and eliminates the otherwise observed effects (e.g., Novemsky et al., 2007). In the absence of such attribution manipulations, however, the fluency-familiarity association affects numerous judgments of everyday importance, including judgments of social consensus, truth, and risk.

As Festinger (1954) noted, we often rely on social consensus information to determine whether an assertion is true or false – if many people believe it, there's probably something to it. One heuristic to estimate social consensus is to assess whether the assertion seems familiar. Accordingly, fluency of processing gives rise to increased estimates of social consensus (Weaver et al., 2007) and facilitates the acceptance of a statement as true (for a review see Schwarz et al., 2007). For example, statements like "Osorno is a city in Chile" are more likely to be judged "true" when they are presented in colors that make them easy rather than difficult to read against the background (Reber and Schwarz, 1999).

Familiarity also figures prominently in intuitive assessments of risk – if a stimulus is familiar and elicits no negative memories, it presumably hasn't hurt us in the past. Accordingly, incidental variables that affect processing fluency also influence peoples'

risk assessments. For example, ostensible food additives are perceived as more hazardous when their names are difficult (e.g., Fluthoractnip) rather than easy (e.g., Magnalroxate) to pronounce (Song and Schwarz, 2009) and stocks with easy to pronounce ticker symbols attract more investors at their initial public offering (Alter and Oppenheimer, 2006). In addition to the mediating role of perceived familiarity observed by Song and Schwarz (2009), intuitive assessments of risk may be further affected by perceivers' positive affective response to fluently processed stimuli (addressed below), consistent with the observation of mood effects on judgment of risk (Johnson and Tversky, 1983) and the beneficial influence of sunny weather on the stock market (Hirshleifer and Schumway, 2003).

Fluency and affect As known since Zajonc's (1968) pioneering mere exposure studies, repeated exposure to an initially neutral stimulus, without any reinforcement, leads to gradual increase in liking. However, repeated exposure is just one of many variables that can increase processing fluency and any other variable that facilitates fluent processing has the same effect. For example, people like the same stimulus more when it is preceded by a visual (Reber et al., 1998) or semantic (Winkielman et al., 2003) prime that facilitates fluent processing, and less when it is preceded by primes that impede fluent processing. In fact, the influence of many variables long known to affect liking and aesthetic preference – from figure-ground contrast to symmetry and prototypicality – can be traced to increased processing fluency (Reber et al., 2004).

This fluency-liking link reflects that fluent processing itself is experienced as pleasant and elicits a positive affective response that [p. 302 ↓] can be captured with psychophysiological measures (Winkielman and Cacioppo, 2001). If this affective response mediates the influence of fluency on liking, it should be eliminated when the positive affect is attributed to an incidental source. Empirically, this is the case as Winkielman and Fazendeiro (reported in Winkielman et al., 2003) demonstrated with misattribution procedures.

Summary and Applications

As the reviewed examples illustrate, people attend to a wide range of feelings as a source of information. However, they are more sensitive to their feelings than to where

these feelings come from and routinely consider incidental feelings relevant to the task at hand. What exactly they conclude from a given feeling depends on the epistemic task they face. Different epistemic tasks bring different lay theories to mind, which link the feeling to the task at hand and serve as inference rules. When feelings are used as a source of information, their use follows the same rules as the use of any other information; hence, the impact of feelings increases with their perceived relevance and applicability and decreases with the consideration of alternative inputs. Whenever people become aware that their feelings may be due to an incidental source, the informational value of the feeling is discredited and people turn to alternative inputs to arrive at a judgment. These regularities hold for moods, emotions, metacognitive experiences, and bodily sensations (Schwarz and Clore, 2007).

The observed use of feelings as a source of information pervades daily life and is not limited to any particular “applied” domain. From the quality of their own lives (Schwarz and Clore, 1983) to consumer products (Isen et al., 1978) and the daily risks they face (Johnson and Tversky, 1983), people arrive at more upbeat evaluations when they are in a positive mood (Schwarz and Clore, 2007) – even when the stakes are high, as illustrated by pervasive mood effects on the stock market (Hirshleifer and Shumway, 2003). Similarly, emotions like anger or fear can shift risk perception and policy preferences (Lerner et al., 2003), as can differences in processing fluency (Song and Schwarz, 2009). Moreover, assessments of truth (Reber and Schwarz, 1999) and the spread of rumors (Schwarz et al., 2007) are profoundly affected by metacognitive experiences, as are consumer (Novemsky et al., 2007) and investment (Alter and Oppenheimer, 2006) decisions. Our feelings are part and parcel of how we think about life and their influence can be observed in any area of investigation.

Feelings and Processing Style

The theory further predicts that feelings or environmental cues that signal a “problematic” situation foster an analytic, bottom-up processing style with considerable attention to detail, whereas feelings or environmental cues that signal a “benign” situation allow for a less effortful, top-down processing style and the exploration of novel (and potentially risky) solutions (Schwarz, 1990, 2002). This does not imply that people in a happy mood, for example, are unable or unwilling to engage in analytic

processing (in contrast to what an earlier version of the theory suggested; Schwarz and Bless, 1991). Instead, it merely implies that happy feelings (and other “benign” signals) do not convey a need to do so; when task demands or current goals require bottom-up processing, happy individuals are able and willing to engage in it. A study that addressed the influence of moods on people's reliance on scripts (Schank and Abelson, 1977) illustrates this point.

Employing a dual-task paradigm, Bless et al. (1996a) had participants listen to a tape-recorded restaurant story that contained script consistent and script inconsistent information. While listening to the story, participants also worked on a concentration test that required detail-oriented processing; in contrast, the restaurant story could be understood by engaging either in script-driven top-down processing or in data-driven bottom-up processing. Happy participants relied on the [p. 303 ↓] script, as indicated by the classic pattern of schema-guided memory: they were likely to recognize previously heard script-inconsistent information, but also showed high rates of intrusion errors in form of erroneous recognition of script-consistent information. Neither of these effects was obtained for sad participants, indicating that they were less likely to draw on the script to begin with. Given that top-down processing is less taxing than bottom-up processing, we may further expect that happy participants' reliance on the script allows them to do better on a secondary task. Confirming this prediction, happy participants outperformed sad participants on the concentration test. In combination, these findings indicate that moods influence the spontaneously adopted processing style under conditions where different processing styles are compatible with the individual's goals and task demands, as was the case for comprehending the restaurant story. Under these conditions, sad individuals are likely to spontaneously adopt a systematic, bottom-up strategy, whereas happy individuals rely on a less effortful top-down strategy. But when task demands (like a concentration test) or explicit instructions (e.g., Bless et al., 1990) require detail-oriented processing, happy individuals are able and willing to live up to the task.

Numerous findings pertaining to a broad range of feelings (from moods and emotions to bodily experiences and processing fluency) and cognitive tasks (from creative and analytic problem solving to persuasion and stereotyping) are consistent with the predictions of feelings-as-information theory (for reviews see Schwarz, 2002; Schwarz

and Clore, 2007). Here I focus on two domains of particular applied interest, namely persuasion and stereotyping.

Applied Implications

Persuasion In general, strong arguments are more persuasive than weak arguments when recipients engage in systematic message elaboration, whereas argument strength exerts little influence when they do not (Petty and Cacioppo, 1986). Accordingly, the impact of argument strength can serve as a diagnostic tool for assessing processing strategy. Studies using this strategy consistently found that happy recipients engage in less, and sad recipients in more, elaboration of counter-attitudinal messages than recipients in a non-manipulated mood (see Schwarz et al., 1991a, for a review). Hence, happy recipients are moderately and equally persuaded by strong as well as weak arguments, whereas sad recipients are strongly persuaded by strong arguments, and not persuaded by weak arguments. Consistent with the feelings-as-information logic, these effects are eliminated when recipients are aware that their mood is due to an unrelated source (Sinclair, et al., 1994). Moreover, the spontaneously adopted processing strategy can be overridden by other variables, such as explicit instructions to pay attention to the arguments (e.g., Bless et al., 1990) or the promise that carefully thinking about the message would make one feel good (e.g., Wegener et al., 1995).

Paralleling the influence of moods, the experience of low processing fluency (which is associated with negative affect and a sense that the processed material is unfamiliar) also fosters the adoption of detail-oriented bottom-up processing, whereas high processing fluency fosters top-down processing (e.g., Song and Schwarz, 2008). Not surprisingly, numerous environmental cues can serve the same informational functions. For example, the same message is less likely to be scrutinized when presented by a communicator with a smiling, happy face than when presented by a communicator with a neutral, somber face (Ottati et al., 1997). Further illustrating the power of contextual cues, Soldat and Sinclair (2001) printed persuasive messages on colored paper. Their recipients were persuaded by strong arguments, but not by weak arguments, when the paper had a depressing blue hue, whereas both types of arguments were similarly persuasive when the paper had an upbeat red hue.

From an applied perspective, these findings suggest that communicators with strong [p. 304 ↓] and compelling arguments have little to gain from putting their audience into a good mood; to the contrary, a somber audience is more likely to elaborate on the substantive implications of the message, facilitating its long-term impact. On the other hand, happy feelings make spontaneous message scrutiny less likely, making smiles, jokes, and upbeat colors promising tools when we have nothing compelling to say.

Stereotyping and impression formation We can form impressions of others by attending to their specific behaviors (bottom-up processing) or by drawing on stereotypic knowledge about social categories (top-down processing). Reiterating the observations from persuasion research, perceivers in a sad mood are more likely to elaborate individuating information about the target person, whereas perceivers in a happy mood are more likely to draw on the person's category membership. This results in more stereotypical judgments under happy than under sad moods (e.g., Bodenhausen et al., 1994; for a review see Bless et al., 1996b). Related research into the influence of brands on product evaluation similarly shows higher reliance on brand information under happy than sad moods (e.g., Adaval, 2001). Note that increased reliance on general knowledge about a brand or group can have positive as well as negative consequences for individual exemplars (be they products or persons), helping exemplars of liked and hurting exemplars of disliked categories. Paralleling the persuasion findings, happy individuals' reliance on category membership information can be overridden by manipulations that increase their processing motivation, such as personal accountability for one's judgment (Bodenhausen, et al., 1994).

Conclusion

In sum, internal and external cues that signal a benign or problematic situation have cognitive and motivational consequences (Schwarz, 2002). Human cognition is tuned to meet situational requirements and problem signals foster vigilance and the adoption of a detail-oriented bottom-up processing style, which is usually adaptive. Signals that characterize the situation as benign, on the other hand, are not by themselves associated with particular processing requirements. They foster reliance on pre-existing knowledge structures and top-down processing, *unless* goals or task demands require otherwise. Which processing strategy facilitates or impedes performance depends on

the specific task. In general, internal or external “problem” signals improve performance on tasks that require analytic reasoning and attention to detail, but impair performance on insight and creativity tasks that require divergent thinking and the exploration of novel solutions (Schwarz, 2002).

Acknowledgments

Completion of this chapter was supported by a fellowship at the Center for Advanced Study in the Behavioral Sciences.

References

- Adaval, R. (2001) Sometimes it just feels right: The differential weighting of affect-consistent and affect-inconsistent product information . *Journal of Consumer Research* , 28, 1–17. <http://dx.doi.org/10.1086/321944>
- Alter, A.L. and Oppenheimer, D.M. (2006) Predicting short-term stock fluctuations by using processing fluency . *Proceedings of the National Academy of Science* , 103, 9369–9372. <http://dx.doi.org/10.1073/pnas.0601071103>
- Barrett, L.F., ed. and Salovey, P. (eds) (2002) *The Wisdom in Feelings* . New York: Guilford.
- Bless, H., Bohner, G., Schwarz, N. and Strack, F. (1990) Mood and persuasion: A cognitive response analysis . *Personality and Social Psychology Bulletin* , 16, 331–345. <http://dx.doi.org/10.1177/0146167290162013>
- Bless, H., Clore, G., Schwarz, N., Golisano, V., Rabe, C. and Wölk, M. (1996a) Mood and the use of scripts: Does being in a happy mood really lead to mindlessness? *Journal of Personality and Social Psychology* , 71, 665–679. <http://dx.doi.org/10.1037/0022-3514.71.4.665>

Bless, H. and Schwarz, N. (1999) Sufficient and necessary conditions in dual process models: The case of mood and processing style . In S. Chaiken, ed. and Y. Trope (eds), Dual-process Models in Social Psychology pp. 423–440. New York: Guilford.

Bless, H., Schwarz, N. and Kemmelmeier, M. (1996b) Mood and stereotyping: The impact of moods on the use of general knowledge structures . In M. Hewstone, ed. and W. Stroebe (eds), European Review of Social Psychology , 7, 63–93. <http://dx.doi.org/10.1080/14792779443000102>

Bodenhausen, G.V. (1993) Emotions, arousal, and stereotype-based discrimination: A heuristic model of affect and stereotyping . In D.M. Mackie, ed. and D.L. Hamilton (eds), Affect, Cognition, and Stereotyping: Interactive Processes in Group Perception , pp. 13–35. San Diego: Academic Press.

Bodenhausen, G.V., Kramer, G.P. and Süsser, K. (1994) Happiness and stereotypic thinking in social judgment . Journal of Personality and Social Psychology , 66, 621–632. <http://dx.doi.org/10.1037/0022-3514.66.4.621>

Bollnow, O.F. (1956) Das Wesen der Stimmungen . Frankfurt: Klostermann.

Bower, G.H. (1981) Mood and memory . American Psychologist , 36, 129–148. <http://dx.doi.org/10.1037/0003-066X.36.2.129>

Clore, G.L. and Byrne, D. (1974) A reinforcement-affect model of attraction . In T.L. Huston (ed.), Foundations of Interpersonal Attraction , pp. 143–170. New York: Academic Press.

Clore, G.L., Schwarz, N. and Conway, M. (1994) Affective causes and consequences of social information processing . In R.S. Wyer, ed. , and T.K. Srull (eds), Handbook of Social Cognition, Vol. 1 , 2nd Edition, pp. 323–418. Hillsdale, NJ: Erlbaum.

Damasio, A.R. (1994) Descartes' Error: Emotion, Reason and the Human Brain . New York: Grosset/Putnam.

Ellsworth, P.C. and Scherer, K.R. (2003) Appraisal processes in emotion . In R.J. Davidson, ed. , K.R. Scherer, ed. , and H.H. Goldsmith (eds), *Handbook of Affective Sciences* , pp. 572–595. New York: Oxford University Press.

Festinger, L. (1954) A theory of social comparison processes . *Human Relations* , 7, 123–146.

Forgas, J.P. (2001) The affect infusion model (AIM): An integrative theory of mood effects on cognition and judgment . In L.L. Martin, ed. and G.L. Clore (eds), *Theories of Mood and Cognition: A User's Guidebook* , pp. 99–134. Mahwah, NJ: Erlbaum.

Friedman, R.S. and Förster, J. (2000) The effects of approach and avoidance motor actions on the elements of creative insight . *Journal of Personality and Social Psychology* , 79, 477–492. <http://dx.doi.org/10.1037/0022-3514.79.4.477>

Frijda, N.H. (1988) The laws of emotion . *American Psychologist* , 43, 349–358. <http://dx.doi.org/10.1037/0003-066X.43.5.349>

Gorn, G.J., Goldberg, M.E. and Basu, K. (1993) Mood, awareness, and product evaluation . *Journal of Consumer Psychology* , 2, 237–256. <http://dx.doi.org/10.1016/S1057-7408%2808%2980016-2>

Greifeneder, R. and Bless, H. (2007) Relying on accessible content versus accessibility experiences: The case of processing capacity . *Social Cognition* , 25, 853–881. <http://dx.doi.org/10.1521/soco.2007.25.6.853>

Greifeneder, R., Müller, P., Stahlberg, D., Van den Bos, K. and Bless, H. (2010) Guiding trustful behavior: The role of accessible content and accessibility experiences . *Journal of Behavioral Decision Making* . DOI:10.1002/bdm.705.

Griffitt, W. and Veitch, R. (1971) Hot and crowded: Influences of population density and temperature on interpersonal behavior . *Journal of Personality and Social Psychology* , 17, 92–98. <http://dx.doi.org/10.1037/h0030458>

Haddock, G., Rothman, A.J., Reber, R. and Schwarz, N. (1999) Forming judgments of attitude certainty, importance, and intensity: The role of subjective

- experiences . *Personality and Social Psychology Bulletin* , 25, 771–782. <http://dx.doi.org/10.1177/0146167299025007001>
- Hansen, J., Dechêne, A. and Wänke, M. (2008) Discrepant fluency increases subjective truth . *Journal of Experimental Social Psychology* , 44, 687–691. <http://dx.doi.org/10.1016/j.jesp.2007.04.005>
- Higgins, E.T. (1998) The aboutness principle: A pervasive influence on human inference . *Social Cognition* , 16, 173–198. <http://dx.doi.org/10.1521/soco.1998.16.1.173>
- Hirshleifer, D. and Shumway, T. (2003) Good day sunshine: Stock returns and the weather . *Journal of Finance* , 58, 1009–1032. <http://dx.doi.org/10.1111/1540-6261.00556>
- Isen, A.M., Shalker, T.E., Clark, M.S. and Karp, L. (1978) Affect, accessibility of material in memory, and behavior: A cognitive loop? *Journal of Personality and Social Psychology* , 36, 1–12. <http://dx.doi.org/10.1037/0022-3514.36.1.1>
- James, W. (1890) Principles of Psychology (Vols. 1 and 2) . New York: Dover. <http://dx.doi.org/10.1037/10538-000>
- Johnson, E. and Tversky, A. (1983) Affect, generalization, and the perception of risk . *Journal of Personality and Social Psychology* , 45, 20–31. <http://dx.doi.org/10.1037/0022-3514.45.1.20>
- Kelley, H.H. (1972) Causal Schemata and the Attribution Process . Morristown, NJ: General Learning Press.
- Keltner, D., Ellsworth, P. and Edwards, K. (1993a) Beyond simple pessimism: Effects of sadness and anger on social perception . *Journal of Personality and Social Psychology* , 64, 740–752. <http://dx.doi.org/10.1037/0022-3514.64.5.740>
- Keltner, D., Locke, K.D. and Audrain, P.C. (1993b) The influence of attributions on the relevance of negative feelings to satisfaction . *Personality and Social Psychology Bulletin* , 19, 21–30. <http://dx.doi.org/10.1177/0146167293191003>

Lerner, J.S. and Keltner, D. (2000) Beyond valence: A model of emotion-specific influences on judgment and choice . *Cognition and Emotion* , 14, 473–493. <http://dx.doi.org/10.1080/026999300402763>

Lerner, S.J., Gonzalez, R.M., Small, D.A. and Fischhoff, B. (2003) Effects of fear and anger on perceived risks of terrorism: A national field experiment . *Psychological Science* , 14, 144–150. <http://dx.doi.org/10.1111/1467-9280.01433>

Martin, L.L., Abend, T., Sedikides, C. and Green, J.D. (1997) How would it feel if...? Mood as input to a role fulfillment evaluation process . *Journal of Personality and Social Psychology* , 73, 242–253. <http://dx.doi.org/10.1037/0022-3514.73.2.242>

Martin, L.L., ed. and Clore, G.L. (eds) (2001) *Theories of Mood and Cognition: A User's Guidebook* . Mahwah, NJ: Erlbaum.

Martin, L.L., Ward, D.W., Achée, J.W. and Wyer, R.S. (1993) Mood as input: People have to interpret the motivational implications of their moods . *Journal of Personality and Social Psychology* , 64, 317–326. <http://dx.doi.org/10.1037/0022-3514.64.3.317>

Matlin, M.W. and Stang, D. (1979) *The Pollyanna Principle: Selectivity in Language, Memory, and Thought* . Cambridge, MA: Shenkman.

Morris, W.N. (1989) *Mood: The Frame of Mind* . New York: Springer Verlag. <http://dx.doi.org/10.1007/978-1-4612-3648-1>

Nisbett, R.E. and Wilson, T.D. (1977) Telling more than we know: Verbal reports on mental processes . *Psychological Review* , 84, 231–59. <http://dx.doi.org/10.1037/0033-295X.84.3.231>

Novemsky, N., Dhar, R., Schwarz, N. and Simonson, I. (2007) Preference fluency in choice . *Journal of Marketing Research* , 44, 347–356. <http://dx.doi.org/10.1509/jmkr.44.3.347>

Ortony, A., Clore, G.L. and Collins, A. (1988) *The Cognitive Structure of Emotions* . New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511571299>

Ottati, V., Terkildsen, N. and Hubbard, C. (1997) Happy faces elicit heuristic processing in a televised impression formation task: A cognitive tuning account . *Personality and Social Psychology Bulletin* , 23, 1144–1156. <http://dx.doi.org/10.1177/01461672972311003>

Ottati, V.C. and Isbell, L.M. (1996) Effects of mood during exposure to target information on subsequently reported judgments: An on-line model of misattribution and correction . *Journal of Personality and Social Psychology* , 71, 39–53. <http://dx.doi.org/10.1037/0022-3514.71.1.39>

Petty, R. and Cacioppo, J. (1986) Communication and Persuasion: Central and Peripheral Routes to Attitude Change . New York: Springer-Verlag.

Pham, M.T. (1998) Representativeness, relevance, and the use of feelings in decision making . *Journal of Consumer Research* , 25, 144. <http://dx.doi.org/10.1086/209532>

Pham, M.T. (2004) The logic of feeling . *Journal of Consumer Psychology* , 14, 360–369. http://dx.doi.org/10.1207/s15327663jcp1404_5

Raghunathan, R. and Pham, M.T. (1999) All negative moods are not equal: Motivational influences of anxiety and sadness on decision making . *Organizational Behavior and Human Decision Processes* , 79, 56–77.

Reber, R. and Schwarz, N. (1999) Effects of perceptual fluency on judgments of truth . *Consciousness and Cognition* , 8, 338–342. <http://dx.doi.org/10.1006/ccog.1999.0386>

Reber, R., Schwarz, N. and Winkielman, P. (2004) Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review* , 8, 364–382. http://dx.doi.org/10.1207/s15327957pspr0804_3

Reber, R., WinkielmanP. and SchwarzN. (1998) Effects of perceptual fluency on affective judgments . *Psychological Science* , 9, 45–48. <http://dx.doi.org/10.1111/1467-9280.00008>

Redlmeier, D.A. and Baxter, S.D. (2009) Rainy weather and medical school admission interviews . *Canadian Medical Association Journal* , 181, 933.

Rothman, A.J. and Schwarz, N. (1998) Constructing perceptions of vulnerability: Personal relevance and the use of experiential information in health judgments . *Personality and Social Psychology Bulletin* , 24, 1053–1064. <http://dx.doi.org/10.1177/01461672982410003>

Ruder, M. and Bless, H. (2003) Mood and the reliance on the ease of retrieval heuristic . *Journal of Personality and Social Psychology* , 85, 20–32. <http://dx.doi.org/10.1037/0022-3514.85.1.20>

Sanna, L. and Schwarz, N. (2004) Integrating temporal biases: The interplay of focal thoughts and accessibility experiences . *Psychological Science* , 17, 474–481. <http://dx.doi.org/10.1111/j.0956-7976.2004.00704.x>

Sanna, L, Schwarz, N. and Small, E. (2002) Accessibility experiences and the hindsight bias: I-knew-it-all-along versus It-could-never-have-happened . *Memory and Cognition* , 30, 1288–1296. <http://dx.doi.org/10.3758/BF03213410>

Savitsky, K., Medvec, V.H., Charlton, A.E. and Gilovich, T. (1998) What, me worry? Arousal, misattribution, and the effect of temporal distance on confidence . *Personality and Social Psychology Bulletin* , 24, 529–536. <http://dx.doi.org/10.1177/0146167298245008>

Schachter, S. and Singer, J.E. (1962) Cognitive, social, and physiological determinants of emotional state . *Psychological Review* , 69, 379–399. <http://dx.doi.org/10.1037/h0046234>

Schank, R. and Abelson, R. (1977) Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures . Hillsdale, NJ: Lawrence Erlbaum.

Schwarz, N. (1987) Stimmung als Information [Mood as information.] Heidelberg: Springer Verlag. <http://dx.doi.org/10.1007/978-3-642-72885-3>

Schwarz, N. (1990) Feelings as information: Informational and motivational functions of affective states . In E.T. Higgins, ed. and R. Sorrentino (eds), *Handbook of Motivation and Cognition: Foundations of Social Behavior* , 2, 527–561. New York: Guilford Press.

Schwarz, N. (1998) Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment . *Personality and Social Psychology Review* , 2, 87–99. http://dx.doi.org/10.1207/s15327957pspr0202_2

Schwarz, N. (2002) Situated cognition and the wisdom of feelings: Cognitive tuning . In L. Feldman Barrett, ed. and P. Salovey (eds), *The Wisdom in Feelings* , pp. 144–166. New York: Guilford.

Schwarz, N. (2004) Meta-cognitive experiences in consumer judgment and decision making . *Journal of Consumer Psychology* , 14, 332–348. http://dx.doi.org/10.1207/s15327663jcp1404_2

Schwarz, N. (2010) Meaning in context: Metacognitive experiences . In L.F. Barrett, ed. , B. Mesquita, ed. and E. Smith (eds), *The Mind in Context* , pp. 105–125. New York: Guilford

Schwarz, N. and Bless, B. (1991) Happy and mindless, but sad and smart? The impact of affective states on analytic reasoning . In J. Forgas (ed.) *Emotion and Social Judgment* , pp. 55–71. Oxford: Pergamon.

Schwarz, N., Bless, H. and Bohner, G. (1991a) Mood and persuasion: Affective states influence the processing of persuasive communications . *Advances in Experimental Social Psychology* , 24, 161–199. <http://dx.doi.org/10.1016/S0065-2601%2808%2960329-9>

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H. and Simons, A. (1991b) Ease of retrieval as information: Another look at the availability heuristic . *Journal of Personality and Social Psychology* , 61, 195–202. <http://dx.doi.org/10.1037/0022-3514.61.2.195>

Schwarz, N. and Clore, G.L. (1983) Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states . *Journal of Personality and Social Psychology* , 45, 513–523. <http://dx.doi.org/10.1037/0022-3514.45.3.513>

Schwarz, N. and Clore, G.L. (2003) Mood as information: 20 years later . *Psychological Inquiry* , 14, 296–303.

Schwarz, N. and Clore, G.L. (2007) Feelings and phenomenal experiences . In A. Kruglanski, ed. and E.T. Higgins (eds), Social Psychology. Handbook of Basic Principles , 2nd Edition, pp. 385–407. New York: Guilford.

Schwarz, N., Sanna, L., Skurnik, I. and Yoon, C. (2007) Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns . Advances in Experimental Social Psychology , 39, 127–161.
<http://dx.doi.org/10.1016/S0065-2601%2806%2939003-X>

Schwarz, N., Servay, W. and Kumpf, M. (1985) Attribution of arousal as a mediator of the effectiveness of fear-arousing communications . Journal of Applied Social Psychology , 15, 74–78. <http://dx.doi.org/10.1111/j.1559-1816.1985.tb02343.x>

Schwarz, N., Strack, F., Kommer, D. and Wagner, D. (1987) Soccer, rooms and the quality of your life: Mood effects on judgments of satisfaction with life in general and with specific life-domains . European Journal of Social Psychology , 17, 69–79. <http://dx.doi.org/10.1002/ejsp.2420170107>

Schwarz, N. and Xu, J. (2010) Constructing heuristics on the spot: Divergent inferences from ease of recall . Unpublished manuscript, University of Michigan.

Sedikides, C. (1995) Central and peripheral self-conceptions are differentially influenced by mood . Journal of Personality and Social Psychology , 69, 759–777. <http://dx.doi.org/10.1037/0022-3514.69.4.759>

Shen, H., Jiang, Y. and Advaval, R. (2010) Contrast and assimilation effects of fluency . Journal of Consumer Research , 36, 876–889. <http://dx.doi.org/10.1086/612425>

Siemer, M. and Reisenzein, R. (1998) Effects of mood on evaluative judgements: Influence of reduced processing capacity and mood salience . Cognition and Emotion , 12, 783–805. <http://dx.doi.org/10.1080/026999398379439>

Sinclair, R.C. (1988) Mood, categorization breadth, and performance appraisal: The effects of order of information acquisition and affective state on halo, accuracy, information retrieval, and evaluations . Organizational Behavior and Human Decision Processes , 42, 22–46. <http://dx.doi.org/10.1016/0749-5978%2888%2990018-0>

Sinclair, R.C., Mark, M.M. and Clore, G.L. (1994) Mood-related persuasion depends on misattributions . *Social Cognition* , 12, 309–326. <http://dx.doi.org/10.1521/soco.1994.12.4.309>

Slovic, P., Finucane, M., Peters, E. and MacGregor, D. G. (2002) The affect heuristic . In T. Gilovich, ed. , D. Griffin, ed. and D. Kahneman (eds), *Heuristics and Biases: The Psychology of Intuitive Judgment* , pp. 397–420. New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511808098.025>

Smith, E.R. and Semin, G.R. (2004) Socially situated cognition: Cognition in its social context . *Advances in Experimental Social Psychology* , 36, 53–117. <http://dx.doi.org/10.1016/S0065-2601%2804%2936002-8>

Soldat, A.S. and Sinclair, R.C. (2001) Colors, smiles, and frowns: External affective cues can directly affect responses to persuasive communications in a mood-like manner without affecting mood . *Social Cognition* , 15, 55–71. <http://dx.doi.org/10.1521/soco.1997.15.1.55>

Song, H. and Schwarz, N. (2008) Fluency and the detection of distortions: Low processing fluency attenuates the Moses illusion . *Social Cognition* , 26, 791–799. <http://dx.doi.org/10.1521/soco.2008.26.6.791>

Song, H. and Schwarz, N. (2009) If it's difficult-to-pronounce, it must be risky: Fluency, familiarity, and risk perception . *Psychological Science* , 20, 135–138. <http://dx.doi.org/10.1111/j.1467-9280.2009.02267.x>

Stepper, S. and StrackF. (1993) Proprioceptive determinants of emotional and nonemotional feelings . *Journal of Personality and Social Psychology* , 64, 211–220. <http://dx.doi.org/10.1037/0022-3514.64.2.211>

Strack, F. and Hannover, B. (1996) Awareness of influence as a precondition for implementing correctional goals . In P.M. Gollwitzer, ed. and J.A. Bargh (eds), *The Psychology of Action* , pp. 579–595. New York: Guilford.

Strack, F., Martin, L.L. and Stepper, S. (1988) Inhibiting and facilitating conditions of the human smile: A non-obtrusive test of the facial feedback

hypothesis . *Journal of Personality and Social Psychology* , 53, 768–777. <http://dx.doi.org/10.1037/0022-3514.54.5.768>

Tversky, A. and Kahneman, D. (1973) Availability: A heuristic for judging frequency and probability . *Cognitive Psychology* , 5, 207–232. <http://dx.doi.org/10.1016/0010-0285%2873%2990033-9>

Wänke, M., Bless, H. and Biller, B. (1996) Subjective experience versus content of information in the construction of attitude judgments . *Personality and Social Psychology Bulletin* , 22, 1105–1113.

Weaver, K., Garcia, S.M., Schwarz, N. and Miller, D.T. (2007) Inferring the popularity of an opinion from its familiarity: A repetitive voice can sound like a chorus . *Journal of Personality and Social Psychology* , 92, 821–833. <http://dx.doi.org/10.1037/0022-3514.92.5.821>

Wegener, D.J., Petty, R.E. and Smith, D.M. (1995) Positive mood can increase or decrease message scrutiny: The hedonic contingency view of mood and message processing . *Journal of Personality and Social Psychology* , 69, 5–15. <http://dx.doi.org/10.1037/0022-3514.69.1.5>

Wegner, D.M. and Vallacher, R.R. (1986) Action identification . In R.M. Sorrentino, ed. and E.T. Higgins (eds), *Handbook of Motivation and Cognition: Foundations of Social Behavior* . New York: Guilford.

Wilson, T.D. and Brekke, N. (1994) Mental contamination and mental correction: Unwanted influences on judgments and evaluations . *Psychological Bulletin* , 116, 117–142. <http://dx.doi.org/10.1037/0033-2909.116.1.117>

Winkielman, P. and Cacioppo, J.T. (2001) Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation leads to positive affect . *Journal of Personality and Social Psychology* , 81, 989–1000. <http://dx.doi.org/10.1037/0022-3514.81.6.989>

Winkielman, P., Schwarz, N., Fazendeiro, T. and Reber, R. (2003) The hedonic marking of processing fluency: Implications for evaluative judgment . In J. Musch, ed. and K.C.

Klauer (eds), The Psychology of Evaluation: Affective Processes in Cognition and Emotion , pp. 189–217. Mahwah, NJ: Erlbaum.

Wyer, R.S. and Carlston, D.E. (1979) Social Cognition, Inference, and Attribution . Hillsdale, NJ: Erlbaum.

Yeh, W. and Barsalou, L.W. (2006) The situated nature of concepts . American Journal of Psychology , 119, 349–384. <http://dx.doi.org/10.2307/20445349>

Zajonc, R.B. (1968) Attitudinal effects of mere exposure . Journal of Personality and Social Psychology: Monograph Supplement , 9, 1–27. <http://dx.doi.org/10.1037/h0025848>

Zajonc, R.B. (1980) Feeling and thinking. Preferences need no inferences . American Psychologist , 35, 151–175. <http://dx.doi.org/10.1037/0003-066X.35.2.151>

Zanna, M.P. and Cooper, J. (1976) Dissonance and the attribution process . In J.H. Harvey, ed. , W.J. Ickes, ed. and R.F. Kidd (eds), New Directions in Attribution Research , 1, 199–217. Hillsdale, NJ: Erlbaum.

Zillman, D. (1978) Attribution and misattribution of excitatory reactions . In J.H. Harvey, ed. , W.I. Ickes, ed. and R.F. Kidd (eds), New Directions in Attribution Research , 2, 335–368. Hillsdale, NJ: Erlbaum.

<http://dx.doi.org/10.4135/9781446249215.n15>

Handbook of Theories of Social Psychology: Volume 1

The Linguistic Category Model

Contributors: Gün R. Semin

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "The Linguistic Category Model"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n16>

Print pages: 309-327

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n16>

[p. 309 ↓]

Chapter 15: The Linguistic Category Model

Gün R. Semin, ed.

Abstract

How did a model of the distinctive terms (predicates) that we use to describe interpersonal events emerge? What were the circumstances that coalesced different inquiries into a model of interpersonal language and how did its applications prove that this was to become an important model of the cognitive properties of language. These are the types of questions that are answered in this chapter, along with how the significance of this model was underlined by its application to how people use language strategically when describing positive and negative behaviors of their ingroups and outgroups. The linguistic category model, which was developed jointly with Klaus Fiedler, has found many applications in answering both fundamental and applied questions. One of the central fundamental issues that this model has revealed is to reformulate the language-cognition interface into a functional: “what is language *for*?” instead of the classic formulation: “what is language?” The answer to the functional question is: “to direct attention.” This formulation opens new vistas for research, such as how people, when pursuing inquiries in question-answer contexts, formulate their choices of predicates that reveal their biases and shape their interviewees’ answers. New directions that can be explored by the linguistic category model are discussed in conclusion.

Introduction

Language is the most pervasive feature of our social environment. It is a tool that we use to implement our goals, intentions, and desires (Semin, 2000a, 2000b). Language extends our capabilities for action irrespective of whether this involves ordering a drink at McDonald's, engaging in a detailed exchange with a sommelier or reviewing a paper, to navigating our ways in a new town with the aid of street names to negotiating our identities in a new relationship. It is undoubtedly the case that there are other incredibly rich nonverbal resources that we recruit consciously and unconsciously in situated interaction (cf. Semin, 2007) but language places us in a world that allows us to travel in time both to the future and the past, to create imaginary worlds that we can share with others, lives that can give us joys and horrors. Language has the potency of ripping us out of our immediate reality and throwing us into different worlds, which can become powerfully possessive realities in their own right.

[p. 310 ↓] It is therefore not surprising that at some point of their life histories people take a reflective stance upon what language is, means, and does to us. Some of us experience the significance of language only when they visit a country and do not understand a single word. Others enter with an intense interest as either avid readers or writers. Some start as chemical engineers, and end up as linguists proposing intriguing hypotheses such as: do our habitual ways of using language shape the way we think, or vice versa (Whorf, 1956).

It is not surprising in my case that language has always constituted a reality that I had to reflect upon because much like the analogy of the fish who realizes the significance of water after landing on the sand, I have been thrown for extended periods from one linguistic community to another and learned that language was important. Indeed, playing with language and writing used to be one of the activities I enjoyed considerably during my high school period. At a very early age, I spent a year in the States and learned English and kept it up fluently after returning to my hometown, Istanbul, because I continuously had British or American friends. I developed a very good working knowledge of French, because my parents spoke fluent French and did so whenever they had a secret to hide from me. They had lots of secrets and I learned fast. I was then sent to the "Deutsche Oberrealschule" in Istanbul and acquired German,

and ended up studying psychology in Germany. After diverse extended stays in different European research centers while firmly anchored on British soil at Sussex University, I arrived in the Netherlands where I acquired Dutch. So, I had to learn to swim in diverse linguistic communities and one would think that that is a good reason to reflect on language. However, I was only interested in language as a tool that had to be mastered and masterly applied in communication until I was well into my PhD research.

It was the research problems I encountered during my thesis research that began to direct me towards language. My PhD was on research that was remotely related to language, namely group decision-making and risk-taking. Trying to understand the dynamics of this question; namely, why do groups make riskier decisions than individuals, I discovered that the question was not quite right. I stumbled upon the semantics of the stimulus items that constituted the subject for the decisions that groups had to make. Indeed, after content-analyzing group discussions and systematically varying different words in the formulation of the dilemmas, it became evident that specific terms in the stimulus items would anchor the focus for discussions. In a dilemma where the group had to advise on whether, for example, an electrical engineer should take a risky job or not, a simple adjective such as “young” was sufficient to focus the discussion and lead to a riskier decision since the group generated a set of beliefs and arguments that were associated with being “young” and were related to risk-taking. If that single word in the dilemma was modified to “old” then the group was likely to generate arguments against taking a risk. These words “primed” the discussion and the direction of the group discussion and decision. This realization, namely that a single word had such a powerful effect in directing the focus of a discussion became important in channeling my attention to the significance of language and raising questions about what role, if any, was played by group processes, since I could also get individuals to generate arguments and produce the same “shifts” (or polarization as it was later labeled) as groups did.

I wanted to learn more about semantics and some of the research I started conducting at the time necessitated this. So, I consulted Gerald Gazdar who was also recently appointed at Sussex University. We had diverse discussions about numerous things, but if my memory serves me well, he recommended not wasting time reading anything on semantics and that I should get on with it on my own, because there were as many semantic theories as there were people writing on it.

[p. 311 ↓] At the time, I took his advice to heart and proceeded with some research which was designed to uncover the degree to which different psychological theories were actually anchored in semantic regularities (e.g., Semin and Chassein, 1985; Semin and Krahé, 1987, 1988). To the extent that the semantic regularities that one discovered mirrored the regularities proposed by theories in personality or otherwise, then the question became: what is the psychological component of the theory? I think that a study I published in 1989 illustrates the nature of my research interest during this period perfectly and reveals that the properties of language can be mistaken for properties of psychological processes. The research I am referring to was a replication of Asch's classic impression formation studies (1946). The method I used to uncover the misattribution of regularities (Semin, 1989) relied on an analysis of a synonym and antonym dictionary and involved computing the semantic overlap between the different stimulus sets (including warm, cold, polite, blunt) and the response scales used by participants in the studies. By developing an index of the strength of semantic associations between the composite stimulus sets and the response items, I was able to show that with the distinctive patterns revealed by actual Asch data could be replicated simply by using a dictionary and no subjects. There was a 90 percent overlap between the dictionary-derived index and the Asch data. This finding suggests that our studies are often examining regularities of the types of tools that we use – particularly language – rather than regularities due to psychological processes.

It was around the time that this research was conducted that my then PhD student, Liam Greenslade, and I (Semin and Greenslade, 1985) became interested in the so-called "systematic distortion hypothesis" by Rick Shweder (e.g., Shweder and D'Andrade, 1980). His contention was that inferences about personality are based on a systematic memory bias whereby co-occurrence frequencies between properties and behaviors of a person are substituted by *similarity* propositions. Liam Greenslade and my take on this supposed general cognitive bias contention for memory-based ratings of persons was that it did not account for functional differences between different types of linguistic forms and that while the argument may hold true for certain linguistic forms it was unlikely to hold true for others. At the time, we made a distinction between *immediate terms* (e.g., verbs, behaviors) and *mediate terms* (e.g., adjectives, traits). This in fact was the turning point in my thinking in that the research we did reflected our realization that different linguistic forms fulfilled different functions in the description

of behaviors and persons. As we showed, while the application of adjectives (mediate terms) was driven by the conceptual relationships (semantics) between the adjectives and were unaffected by situated features that were also relevant, this was not the case for verbs (immediate terms). The use and application of verbs displayed considerable systematicity as a function of the situated features in which people found themselves. These were sensitive to the contingencies of the environment unlike the proposition advanced by the systematic distortion hypothesis.

This research (Semin and Greenslade, 1985) constituted a turning point for me. I moved from my focus on investigating whether assumed psychological processes were merely reproductions of semantic relationships found in everyday language to developing an increased interest in the meta-semantic properties of language. The person who was responsible in accelerating this perspective was my friend and colleague Klaus Fiedler with whom we had extended discussions on the subject. The outcome of different ideas turned out to be the paper, which is the metasemantic model of interpersonal language, namely the linguistic category model (LCM). Our discussions focused on specifying the crude taxonomy that Liam Greenslade and I had advanced by, first, limiting the domain to interpersonal language and, second, refining what we had termed "immediate terms." Around this time, Brown and [p. 312 ↓] Fish's (1983) fascinating paper on the causality implicit in language had also appeared and a visiting American colleague with whom we had discussed our ideas drew our attention to it. Reading this paper influenced our thinking and shaped the way we developed our taxonomy of interpersonal language.

We developed our ideas to investigate the linguistic devices that are used in the description of social events and persons – the actors in such situations. The question was: What are the distinctive properties that such devices have? This resulted in what is now known as the LCM (Semin and Fiedler, 1988, 1991) which is a classificatory approach to the domain of interpersonal language and consists of interpersonal (transitive) verbs that are tools used to describe *actions* (help, punch, cheat, surprise) or psychological *states* (love, hate, abhor) and *adjectives* and *nouns* that are employed to characterize persons (e.g., extroverted, helpful, religious). The question addressed by the LCM is: what are the distinctive interpersonal lexical units that are available and are there systematic cognitive properties that they share in differing degrees?

The idea of examining language as possessing cognitive properties in the late 1980s was an unusual one and indeed the publication of the first paper on the LCM took an extensive reviewing and revision process of over two years until it was accepted for publication by the then editor-in-chief of the *Journal of Personality and Social Psychology*, Jim Sherman. The fortunate aspect of this extended scholarly discussion was that the reviewers and editor were constructively involved in drawing attention to ways of improving the taxonomy, its criteria, and the clarity of the presentation. We were very fortunate in Jim Sherman's informed selection of reviewers, of whom one – we think – was Roger Brown, and the other was Bill McGuire. There were other constructive reviewers but our evidence was not strong enough to infer their identity. And, yet at the time when the information processing based social cognition tradition was at its height, presenting a paper about the cognitive properties of language was unusual. So, let me now turn to the LCM to detail the nature of this taxonomy, which was driven by an idea that language is a tool.

The Linguistic Category Model

This model of interpersonal language furnishes the means by which it is possible to identify the nuances of how people use interpersonal terms when they are representing social events in communication, and thus is informative about how verbal behavior is driven strategically by psychological processes and communication constraints. This is made possible by providing a systematic model of the linguistic terms (verbs, adjectives, and nouns) that we use in communicating about social events and their actors.

In this model, a distinction is made between five different categories of interpersonal terms, namely Descriptive Action Verbs (DAV), Interpretative Action Verbs (IAV), State Action Verbs (SAV), State Verbs (SV), and Adjectives (ADJ) (cf. Semin and Fiedler, 1991). The distinction between the categories is obtained on the basis of a number of converging grammatical tests and semantic contrasts (cf. Bendix, 1966; Brown and Fish, 1983; Miller and Johnson-Laird, 1976).

DAVs are the most concrete terms and are used to convey a description of a single, observable event and preserve the perceptual features of the event (e.g., "A punches B" whereby punching is always achieved by means of a fist). The second category_(IAVs)

describes specific observable events. However, these verbs are more abstract in that they refer to a general class of behaviors and do not preserve the perceptual features of an action (e.g., “A hurts B”). SAVs and SVs refer to psychological states while DAVs and IAVs do not. SAVs refer to the affective consequences of actions that are not specified any further (to amaze, surprise, bore, thrill, etc.) but can be supplied when asked (e.g., “Why was she surprised?”). SVs typically describe an unobservable affective or mental [p. 313 ↓] state and not a specific event (e.g., “A hates B”). Finally, adjectives (e.g., “A is aggressive”) constitute the most abstract category. These generalize across specific events and objects and describe only the subject. They show a low contextual dependence and a high conceptual interdependence in their use. In other words, the use of adjectives is governed by abstract, semantic relations rather than by the contingencies of contextual factors. The opposite is true for action verbs (e.g., Semin and Fiedler, 1988; Semin and Greenslade, 1985). The most concrete terms retain a reference to the contextual and situated features of an event. This dimension of abstractness-concreteness of interpersonal predicates has been operationalized (e.g., Semin and Fiedler, 1988) in terms of a number of different inferential features or properties with DAVs (hit, kiss) being the most concrete category. IAVs (help, cheat) are more abstract. SVs (like, abhor) follow next, and ADJs (friendly, helpful) are the most abstract predicates.

It is important to note that the properties by which abstractness-concreteness had been operationalized are *generic to the entire predicate classes* represented in the LCM. The types of meanings or implications as defined by the distinctive inferential properties of the LCM are different from the more conventional study of meaning, namely semantics. The more conventional approaches in linguistics are the study of meaning in terms of *semantic fields*, *semantic relations*, or the analysis of lexical items in terms of *semantic features* to investigate the semantic component of a grammar's organization. One may refer to the meaning domain identified by the LCM as meta-semantic since the inferential properties apply across semantic fields and are also distinctive in that they escape conscious access (Franco and Maass, 1996, 1999; Semin, 2006; Von Hippel et al., 1997).

Lcm: The Early Applications

The development of the LCM was followed by a rich series of experimental studies which were all concerned with examining systematic use of the different predicate categories and had an eye to either explaining existing phenomena as a consequence of strategic language use (e.g., actor-observer discrepancy and fundamental attribution error, Semin and Fiedler, 1989), or introducing new phenomena that resulted from systematic differences in language use. A superb example for the latter is the research field initiated by Anne Maass on strategic language use by members of in- and outgroups (Maass, 1999; Maass et al., 1989).

Design Processes for Stereotypes

What are the “design processes” involved in communicating about stereotypes? Which lexical categories of the LCM do we choose when we are conveying events involving members of our in- or outgroups? This was the issue that was addressed by Anne Maass (1999 for a review) and her colleagues. Stereotypes are the emergent results of socially situated interactions between individuals, rather than a product that resides within the head of an individual, an idea whose roots can be dated in Lippman's (1922) famous “pictures in the head” metaphor. The research on strategic language use and stereotypes, initiated by Anne Maass and her colleagues, on how language is used as a tool reveals that the way we represent events involving in- and outgroups is a dynamic design process driven by cognitive and motivational factors.

The linguistic intergroup bias (LIB, Maass and Arcuri, 1992; Maass et al., 1989, 1995) involves a linguistic strategy for individuals to describe positive ingroup and negative outgroup behaviors in relatively abstract terms (e.g., adjectives), implying that the behavior in question is due to enduring dispositions or the actor's stable characteristics. Conversely, negative ingroup and positive outgroup behaviors are typically described in relatively concrete terms, implying the incidental or situational specificity of the [p. 314 ↓] behavior, and hence an external attribution of the behavior.

Two mechanisms have been postulated to account for these systematic differences that people display when they are talking about positive and negative out- and ingroup behaviors. One possible mechanism of the LIB could be motivational (Maass, 1999; Maass et al., 1989) having to do with the fact that abstract descriptions of positive ingroup behaviors and of negative outgroup behaviors portray the ingroup in favorable and the outgroup in unfavorable terms implying that these behaviors are due to enduring characteristics. Similarly, concrete depictions of negative ingroup behaviors minimize their significance as evidence for corresponding group characteristics, as do concrete depictions of positive outgroup behaviors. In other words, those linguistic strategies serve to protect the perception that the ingroup is superior to the outgroup. Another mechanism to account for these patterns of language use is provided by a cognitive or expectancy account which states that expected behaviors are described with abstract language and unexpected behaviors by the use of concrete predicates (e.g., Rubini and Semin, 1994). Both processes appear to be operative depending on the motivational circumstances under which the strategic language is produced (Maass, 1999; Maass et al., 1995). The expectancy driven mechanism has been termed the linguistic expectancy bias (LEB; Wigboldus et al., 2000).

The extensive research on the LIB/LEB extends the language social cognition interface, by showing precisely how both cognitive and motivational processes systematically influence language as a tool that renders these processes as action. Indeed, subsequent research investigating the impact of these messages upon third parties has shown that third parties draw the implications for which the messages are designed (e.g., Werkman et al., 1999; Wigboldus et al., 2000). Extending this line of thinking, and building on their earlier work (2003), Douglas and Sutton have shown that linguistic choices implicitly convey to listeners of messages the type of relationship that holds between the producers of such messages and the targets of their message (Douglas and Sutton, 2006).

The communication cycle between psychological processes, message production, and message impact is closed in a series of studies reported by Reitsma-van Rooijen, et al. (2007). These authors have shown that when participants receive messages that are either concretely or abstractly formulated about their own positive and negative behaviors (e.g., success or failure) from another person, then this affects their perceived social distance to the sender of the message. Similar to the findings reported by Rubini

and Kruglanski (1997) in an entirely different context, abstract messages about positive behaviors and concrete messages about negative behaviors produced perceived social proximity. In contrast, receiving concrete messages about positive behaviors and abstract messages about negative behaviors led to judgments of social distance to the sender of the message.

There are different ways of looking at these studies that reveal strategic language use in intergroup and interpersonal contexts. These studies take the language-cognition interface outside of the boundaries of the individual “mind” and investigate the effects of psychological processes upon strategic language use in the first instance. How do more broadly conceived cognitive and motivational processes influence the “design process” people use when producing a message or utterances? The next step is to investigate the impact of such messages. This can be done in a number of ways as mentioned earlier. How do neutral participants decipher what is in a message? How do involved parties whose doings are talked about react to such senders? More importantly, how does language use contribute to the regulation of social behavior?

The language-cognition (or rather psychological processes) question acquires an entirely different complexion as can be seen [p. 315 ↓] in the studies I summarized above. The question becomes an investigation of how psychological processes contribute to the design of language use, and consequently how different utterances as the result of design processes direct attention to different facets of the same event, thereby influencing the way a receiver perceives the same event and how such a process contributes to the regulation of social interaction. This constitutes a major shift in the way the language question has been posed in psychology. Instead of posing the question “What is language?” and examining the language psychological process interface as interconnected interpsychological inquiry, the question becomes, “What is language for?,” and transforms our approach to language as a tool that is used to extend cognition into communication. This is the transition that the research on the linguistic intergroup bias and other related research with the LCM introduced in my thinking.

Thus, the central mission that emerged from this research led to a reformulation of the classic language-cognition interface question that has been with us for a very long time. The contribution of a functional take on language, namely, “What is language

for?,” leads to a translation of the very same classic puzzles into contextually situated problems that open increasingly inclusive research challenges. The central mission that emerges is interfacing psychological processes within a communication context with a view towards developing a better understanding of the regulation of social behavior. Thus, before proceeding with how the work on the LCM transformed my thinking and the future output from this model, I shall detail the distinction between the classic and functional approach towards understanding the interface between language and cognition.

What is Language (for)?

There are many questions that one can pose to address the different puzzles that language presents, one of the most distinctive human artifacts. Not surprisingly, the question one asks shapes the nature of the answer, and nearly without exception, this *a priori* determination of an answer by the question escapes our awareness. Every question is the result of a set of presuppositions, and it is no different when questions about language are formulated. Very subtle differences may have substantial consequences. We can ask what language *is*, or alternatively, what language is *for*. The answer to the former would treat language as an entity with an existence independent of its “natural habitat” (communication), namely an object that is timeless, subjectless, and with an internal logic. The latter formulation, which introduces an apparently minor modification, namely what is language *for*, all of a sudden changes the nature of our inquiry dramatically, and our interest shifts substantially from an inquiry about a decontextualized abstract entity to a live and active instrument serving diverse purposes in a dynamic interpersonal context. Indeed, these two questions about language represent the classic and the more recent approaches towards understanding language and their psychological implications. In the following, I shall present these two perspectives in turn and conclude with some issues about the implications arising from the second question.

What is Language?

“What is language?” typifies the traditional approach to language – a view that still informs much research in social cognition. In this view, language is a disembodied structure and dissociated from real time. Language is regarded as a set of symbols and “rules” about how to combine them and both are “virtual and outside of time” (Ricoeur, 1955). Language is considered in the abstract, and “without a subject.” This lens presents language as an extra-individual and systematic set of abstract properties and a life of its own. Consequently, it is “subjectless,” “timeless,” and disembodied. This assumption enables the examination of language in terms of the relationships between its specific properties [p. 316 ↓] in detached abstraction (e.g., lexical semantics, grammatical categories, word order).

Such an approach to language does not lend itself readily for a *social* examination of how language interfaces with psychological processes, but it does pair with a view of cognition, which treats cognition in a manner comparable to the way language is treated in the “What is language?” perspective. In the classic (and largely current) approach, cognitive processes are also thought of in a disembodied, timeless, and subjectless manner. In this view, both language and cognition are separated from real time.

Detaching both language and cognition from situated action has led to a perspective where both are located in and examined from an *individual perspective*. The central question then becomes the mutual influence between language and social cognition. This mutual influence is understood in terms of two sets of “inner representational” systems, rather than interaction in a social world (e.g., dialogue, etc.). Given these assumptions, the interface between language and social cognition becomes an investigation of the interaction between two different *representational* modules in an intrapersonal “forum,” with social cognition referring to individual processes – encoding, representing, thinking, retrieving, and so on. The classic and ongoing debate on linguistic relativity (Whorf, 1956) in psychology is located precisely at the center of this classic viewpoint. The research agenda was conceived by Lenneberg and his colleagues (Brown and Lenneberg, 1954; Lenneberg, 1953; Lenneberg and Roberts, 1956). The central question: *Does the structure of a given language affect the thoughts (or thought potential), the memory, the perception, the learning ability of those who*

speak that language? (Lenneberg, 1953: 463, italics here). This early crystallization of how to tackle the topic defined and determined the research focus on the interface between language and social cognition, and it is current in both psychology and social psychology: *Do linguistic structures influence nonlinguistic categorization, memory, perception, attention, thinking, and so on?* The research questions are seated in the “mind” and the language and social cognition relationship becomes an analysis of the construction and manipulation of inner representations. The more recent versions of the view also share the same individual centered representational perspective (e.g., Hardin and Banaji, 1993; Hoffman et al., 1986; Hunt and Agnoli, 1991). The research question can be illustrated with the following example. If two cultures code the color spectrum differently in their respective linguistic communities, do they then perceive and represent colors incommensurably or not (cf. Özgen, 2004)? Not surprisingly, thinking of language and cognition in this way leads to the fascinating and classic issues that have occupied many scholars about the relationship between language and cognition, and their mutual influence (cf. Whorf, 1956).

What is Language? Language Locked in the Mind

An interesting instance for the case of “What is language?” was provided by a “phenomenon” that has attracted considerable research interest in social cognition and termed by Brown and Fish (1983) “the *psychological causality implicit in language*.” As I shall note shortly, this is a much broader field than the focus that Brown and Fish have introduced. The question that the research in this field attempted to answer was: What types of systematic “inferences” do transitive verbs that describe interpersonal events (e.g., to help, to cheat, to respect, to surprise) convey?

The research emerging from different strands has focused on the inferences that people make as a function of the verb type in minimal sentences (e.g., “subject-verb-object” sentences such as “John helps David” or “John likes David”). The difference between the two example sentences is the verbs the first referring to actions (e.g., hit, help, cheat) and the second to states (e.g., like, respect, hate). This research has revealed a multitude of different types of inferences [p. 317 ↓] over a period if more

than 40 years (see, for detail, Semin, 2007). Indeed, the origin of this question can be found in a research program initiated by Abelson and his collaborators (e.g., Gilson and Abelson, 1965, Abelson and Kanouse, 1966) much before Brown and Fish's work (1983). The phenomenon was investigated independently in linguistics with a "pronoun disambiguation" paradigm (e.g., Caramazza et al., 1977; Ehrlich, 1980; Garvey and Caramazza, 1974; Garvey et al., 1976). Another perspective focused on the role that interpersonal verbs play in *text comprehension* (e.g., Garnham and Oakhill, 1985; Garnham et al., 1992, 1996; Greene and McKoon, 1994; McKoon et al., 1993).

The interest in interpersonal verbs has focused on the systematic and different inferences mediated by *action* and *state verbs* has revealed a range of inferences varying from the assignment of cause or agency to the event (e.g., Brown and Fish, 1983), different inductive-deductive inferences (e.g., Gilson and Abelson, 1965), implicit quantifiers (Kanouse, 1972), differential salience of sentence subject and object (Kasoff and Lee, 1993), differential disambiguations of personal pronouns (e.g. Garvey et al., 1976; Garvey and Caramazza, 1974), different foci of what preceded and antecedent the event in the stimulus sentence (Au, 1986; Fiedler and Semin, 1989), *inter alia*. The list is long (cf. Semin, 2007 for an overview).

The nature of these findings point to a set of paradoxes that emerge if one approaches the language-cognition interface in this domain with a "What is language?" perspective. The sheer fact of the great number of demonstrated systematic inferences mediated by interpersonal verbs poses a number of problems. One of these has to do with the assumptions about the supposed psychological processes (inferences) that are "activated." If all of the recorded systematic inferences were automatic and autonomous as they are assumed to be (e.g., Brown and Fish), and were to be coactivated, then the sheer cognitive load that this would introduce would in all likelihood bring the "cognitive apparatus" to a grinding halt and lead to a rather pathological state in the subject. This problem arises due to the assumption that both language and cognition are intrapersonal processes whereby a function perspective, and consequently the adaptive action implications of both language and cognition, is absent. This is not to deny the empirical reality of the findings, but to question their conceptual grounding. The more interesting question that arises from a situated and functional perspective is what the multitude of systematic inferences mean.

The tacit treatment of language and cognition as inner representational systems rather than being “for” adaptive interaction with the world renders the puzzle of inferential processes an individual centered one. Language and cognition are assumed to “happen” *within an individual*. Language and cognitive processes remain disembodied, timeless, and subjectless. These findings do not make much sense if seen from a traditional perspective on language and cognition, which is not informed about a communicative or interpersonal context, and the chief function that language serves.

The paradoxical problems that this type of conceptual framing faces arises as a consequence of the traditional or geocentric view that dismisses the inherently variable processes that characterize the adaptive flexibility of cognition as well as language as a property that emerges in the interaction between the organism a social environment (Smith and Semin, 2004), and the function of language as a tool in extending cognition in real or imagined communicative contexts (cf. Semin, 2000b, 2001, 2007). As I have elaborated elsewhere (Semin, 1998, 2007) the varied experimental scenarios in which the different inferences are revealed constitute different situations where the dependent variables (about causal agency, inductive or deductive inferences, etc.) constitute different situated contexts that demonstrate the malleability of language rather than some underlying psychological process driving different types of inferences. To understand what this means [p. 318 ↓] more specifically, we have to adopt a different more heliocentric view of language. This invites the question: What is language for?

What is Language for? Language as a Tool for Action

Posing the question, “What is language for?,” introduces an alternative way of looking at this relationship, namely a functional view by considering the language-cognition interface in a *language use context*. In this functional view, language is treated as a *tool* and a means to extend cognition in the implementation of action (e.g., Semin, 2000a, 2001). From this perspective, language is for use, and in general terms language use is a *design process* that extends cognitive and motivation processes of a speaker with a view to direct the attention of a listener to some aspect of social,

physical, or psychological reality. In other words, language is used in a communicative context with a view to structure the cognition of an addressee by driving his or her attention. Obviously, this is an interactive process and not unidirectional. Seen this way, cognition can refer to both (1) processes which contribute to how a speaker shapes a communicative act (production processes), (2) to those processes that contribute to how a communicative act (a message) is received by an addressee (comprehension processes), and (3) the entirety of communication itself, namely independent of the individual productions, as a regulator of joint action (see Hutchins, 1996).

This way of conceptualizing language introduces a way of thinking about language that is somewhat alien to how it is conceptualized traditionally in psychology. It effectively means that *language* – as a tool – *has cognitive properties*. Treating language as a tool (Semin, 1998) means literally that: different linguistic units and their conventionalized combinatorial possibilities constitute devices to do things. The distinctive characteristic that language as a tool shares with devices in general is that tools have evolved to meet the dual demands involved in solving a problem by adapting to the features or properties of a particular task (e.g., cutting a piece of paper) as well as the constitutional features of human capabilities and body (e.g., the shape of the hand). This *dual adaptation* gives us capabilities or powers to do things that we do not have by nature. For instance, a pair of scissors consists of two blades with a ring shaped handle, which is adapted to the features of a handgrip. These blades are so pivoted that their sharp edges move one against the other, which is the perfect adaptation to perform a neat cutting operation on paper or materials. The distinctive feature of any tool, as in the case of a pair of scissors, is that *knowledge is downloaded onto it*. The way it has been engineered, the sharp blades, the way the blades are pivoted, the rings on the blades constitute a remarkable achievement in terms of facilitating the interface or the coupling between the human constitution and the human goal. The distinctive feature of any device is that it is embodied (Semin and Smith, 2008). Tools are shaped by the constraints of human biological constitution on one hand and the types of adaptation that the tasks at hand require on the other. This way of looking at language takes it from the exclusively representational perspective and grounds it into an embodied framework. Language as a *tool* contains functional knowledge and the type of “task” faced in a communication context introduces a “design process,” which results in unique communicative acts to resolve the task. One particular implication of this view

is to examine the types of tools and their characteristics in terms of the knowledge downloaded onto them.

What is Language for? Driving Attention!

Any social event can be represented in a number of different ways with a number of different linguistic devices that give different nuances to the aspect of the reality to which we would like to draw the listener's attention. Concretely, this means to draw attention to specific aspects of the physical, psychological, and social environment in communication (e.g., an event whereby David's fist [p. 319 ↓] connects with John's jaw). To do so, the speaker has a variety of options in terms of the devices that she can use to draw attention to the very same reality (e.g., He punched John; He hurt John; He hates John; He is aggressive). Particular psychological processes and the social context drive the choice of specific lexical categories and contribute to the design of communicative acts (utterances) directing attention to different aspects of reality. Accordingly, different linguistic devices (e.g., punch, hurt, hate, aggressive) serve different perceptual functions.

This particular take on the relationship between language and cognition assigns different roles upon language and cognition. Language becomes the tool by means of which cognition is implemented in action whereby language functions as an attention-driving device in an interpersonal context.

I shall now turn to another application of the LCM – to illustrate the “What is language for?” perspective. This application is the examination of how “interpersonal verbs” can be used in the formulation of questions and how they shape answers by drawing attention to different thematic foci and contribute to the shape of inferences.

Lcm: Applications in Interview Situations – the Question-Answer Paradigm

What are the implications of a “What is language for?” perspective for the paradoxical literature on interpersonal verbs referred to above? What is the consequence of taking interpersonal verbs out of representational contexts as mentioned earlier, and inserting them into an experimentally induced dialogical context? Dialogical contexts are situated and thus driven by specific goals or tasks. Take a specific event, such as witnessing a “sailor at La Scala to see a soprano.” One can *direct a listener’s attention* to different facets of this event by choosing one’s verbs as well as who occupies the sentence subject or object. For instance, we may say “Why does the sailor *visit* the soprano?,” or “Why does the sailor *like* the soprano?,” thereby changing the focus of attention to the sentence subject or object depending on the respective sentences used. Saliency, but also respectively agency, for the event is thus modified by the choice of verb in the question and indeed it has been shown that answers to such questions focus respectively on the sentence subject (question with action verb) or object (question with state verb) (Semin, 2000b; Semin et al., 1995). While the sailor is in the semantic role of agent in the former, he occupies the semantic role of experiencer in the second sentence. Alternatively, we can formulate questions that have been used in prior research. Posing the question: “How likely is it that sailors like sopranos?” is inductive generalization. The question “How likely is it that the sailor will like the Maria Callas?” is deductive generalization, and so on. Thus, given the very same event (stimulus sentence), the formulation of the sentence (in these particular instances the question) is used as an attention-driving device. This formulation “compels” the respondent with the implicitly operating rules of conversation (Grice, 1975) and obliges her to follow the social contract that is issued – thus having to produce the “appropriate” answer. The appropriate answer is driven by the conventional cognitive properties of interpersonal verbs (e.g., Semin and Fiedler, 1988) and these are not determinate because language is not a determinate tool. As von Humboldt observed, language “makes infinite use of finite media” (1836: 70) whose “synthesis creates something that is not present *per se* in any of the associated constituents” (1836: 67). The point of this exercise is to simply demonstrate that the type of question focuses attention on a distinctive feature

or aspect of the social event observed and drives the answer of the recipient of the question.

While the above considerations remain thought experiments, there is a collection of research that has investigated the attention-driving or thematic-focusing function of interpersonal verbs (action and state verbs) in [p. 320 ↓] experimentally situated contexts. The two types of interpersonal verbs function as structural devices that differentially focus attention (Stapel and Semin, *in press*) upon either the sentence subject or object in the question. This function of interpersonal verbs is what has been termed by some (e.g., Kay, 1996) “perspectivization” or “topicalization” (Fillmore, 1968). This attention-driving function of interpersonal verbs and their psychological implications has been examined in an experimentally induced communication paradigm termed the “question-answer paradigm” (QAP; Semin, 2000b). This paradigm was introduced by Semin et al. (1995) who systematically manipulated the verb in question formulation and examined whether these two verb types (action versus state verbs) systematically bias a target's answer as well as the implications of this answer (cf. De Poot and Semin, 1995; Rubini and Kruglanski, 1997; Semin and De Poot, 1997a, 1997b). The research by Semin et al. (1995) indicated that questions formulated with action verbs (e.g., “to help,” “to write”) cue the logical *subject* of a question as the causal origin of answers. Questions formulated with state verbs (e.g. “to love” or “to like”) cue the logical *object* of a question as the causal origin for answers. Thus, if asked such simple and mundane question as, “Why do you *own* a dog?” (using an action verb) participants are prompted to respond by referring to themselves (the subject of the question) as the causal agent in the answer; for example, by stating, “Because *I* enjoy the companionship that dogs provide.” If one is asked “Why do you *like* dogs?” however, one is prompted to respond by referring to the object itself, for example, “Because *dogs* are good companions” (e.g., Semin and de Poot, 2007a). Furthermore, the abstractness level of questions was also shown to influence the abstractness level of the answers (cf. Semin, 2000b, for a review). Thus, more abstractly formulated questions tend to elicit more abstract answers. Interestingly, Rubini and Kruglanski (1997) set out to investigate the further implications of such differences in abstractness not only for how such verbs steer thematic focus but also the implications of dialogue sessions controlled over experiments for perceived interpersonal proximity and distance. They did so by examining if individuals had the impression that they disclosed more about

themselves when asked questions formulated with action-verbs (concrete) as opposed to state-verbs (abstract) as well as examining the moderating role of need for closure. In their first experiment designed to investigate these issues, Rubini and Kruglanski (1997) had participants under high (versus low) need for closure (operationalized via ambient noise) rank order questions out of a list in terms of their likelihood of using them in a real interview. The list included 32 questions, eight questions on each of four different topics. It was found that participants under noise (versus no noise) assigned higher ranks to questions characterized by higher (versus lower) level of abstractness. In a follow-up study, questions selected by participants under high (versus low) need for closure were found to elicit more abstract answers from respondents, and ones focused more on the logical *object* (versus *subject*) of the question. In addition, respondents reported that they felt less friendly toward the interviewer whose questions were more (versus less) abstract. Finally, in a third study the results of the previous two experiments were replicated in a free-interaction context. Interviewers with high (versus low) need for closure asked more abstract questions, which in turn elicited more abstract answers and ones focused more on the logical *object* (versus *subject*) of the question, and elicited lesser friendliness from the interviewee. These results suggest that the permanence tendency induced by the need for nonspecific closure may affect the level of linguistic abstractness, and in so doing may imbue the nascent social relations among conversation partners.

In this section, I reviewed some of the research that reveals the attention driving properties of interpersonal verbs in terms of both the types of answers they induce when [p. 321 ↓] used in question formulation as well as their psychological consequences for interpersonal relationships. This stands in contrast to the traditional or geocentric approach in which the interpersonal context of language is absent. Admittedly, the traditional approach has yielded a wealth of findings and yet the very nature of the findings gives rise to internal inconsistencies and paradoxes as I noted earlier.

Conclusion

The chief point that emerged in the pursuit of the different ways in which the LCM can serve research was to shape a different way of formulating the relationship between

language and psychological processes in general and cognition in particular. The driving idea that emerged in the course of the research that I and others engaged in was to understand the issues that connect language and human function in terms of: “What is language for?,” namely its function. This perspective was developed from the following arguments.

- Language is a tool.
- As any tool it is characterized by dual adaptation, namely to (1) the specific communicative tasks at hand, and (2) the human constitution.
- As a consequence of dual adaptation, language provides an ecological niche that contains information.
- The main function of language is to implement psychological processes (cognitive and motivational) in action.
- The composition of utterances (Semin, 2006) is the result of a “design process,” which is driven by psychological processes, situated demands, and conversational rules.
- The main function of language use is driving attention.

In closing this contribution, I shall elaborate briefly upon the last argument above, which is probably the most important argument deriving from the perspective adopted on language and psychological processes. In a sense, the last argument may be self-evident, particularly since everyday examples of strategic use of language, in other words *political spin*, have become an integral part of our folklore. Political spin is in principle the strategic use of language in order to attract attention or detract attention from certain facts of life. Let me briefly illustrate. One can label specific people as *freedom fighters* and thereby direct attention to specific aspects of these people including sowing the associated seeds of images that have to do with virtue, self-sacrifice, patriotism, which help go beyond the “label” given – to paraphrase Jerome Bruner. By labeling the very same people *terrorists* we conjure entirely different images and associations (cruel fanatics, possessed psychopaths, etc.). This is probably the obvious attention driving function of politically driven design processes and one that operates at a semantic level.

At this semantic level of analysis, the attention of the receiver is driven by specific features and properties associated with a category label. This is perhaps an obvious

point from the traditional approach to language (What is language?). Language directs people's attention and perceptual focus, and different linguistic devices direct attention to different aspects of reality. This is an idea that is also at the core of Whorf's (1956: 221) linguistic relativity hypothesis, which suggests that the use different "grammars" direct people to different types of observations and evaluation of events. But what is distinctive about this assumption is that the focus is upon the content of attention driving which is more about associations and specific topics, themes, or beliefs. Moreover, the entire process is supposed to be played out in a person's head. This is a *surface-level process* driven by lexical meaning or meaning as moderated by the "rules" governing the permissible word compositions in utterances. This is in fact the way the classic work on the language-cognition interface has proceeded and produced considerable insights by comparative studies across linguistic communities. Examples are abundant, such as the types of lexical categories [p. 322 ↓] that are available to describe persons in English and Chinese (e.g., Hoffman et al., 1986); the domain of language responsible for spatial locations (Majid et al., 2004), variations in the availability of basic color terms across a diversity of linguistic communities (e.g., Berlin and Kay, 1969); differences between languages in gender marking (e.g., English versus Turkish) or pronoun drop (e.g., Kashima and Kashima, 2003), or lexical categorizations emotional states (Semin et al., 2002). Notably, all these cultural differences are domain specific and can be seen as culturally marked manners by which attention is driven to specific aspects of the social, physical, and psychological reality.

The pursuit of the implications furnished by the LCM drove me to a different perspective on this research. If language is an attention-driving device, then it may be the case that specific linguistic devices, such as the metasemantic properties of the linguistic categories identified by the LCM (Semin and Fiedler, 1988) are themselves functionally organized to drive attention in a generic fashion. This leap, although implicitly if not explicitly always inherent in the conceptualization of the LCM, was never the direct object of empirical investigation

Diederik Stapel and I, after lengthy discussions over dinner in Amsterdam, used the LCM as a conceptual framework to investigate the hypothesis that different linguistic devices within the same language may have generic, metasemantic effects on cognition (Stapel and Semin, 2007). In our studies, we focused on the abstractness-concreteness dimension of the LCM, and suggested that if concrete terms such as action verbs (to

kiss, to hit, to push) are used predominantly in situated contexts and refer to the specific details of a social event, then their obvious function – aside from providing a semantic representation of the event – is to *draw attention to the situated, local features of the event* (Stapel and Semin, 2007). At the other end, abstract terms such as adjectives detract from any transient situated features of an event and direct global focus (John is aggressive).

From this it follows that different predicate classes are likely to direct attention to different features of an object. Now, the research question is not about specific semantic categories (e.g., freedom fighter) and their semantic associates. It is about metasemantic features of interpersonal language (Semin, 2000). We (Stapel and Semin, 2007) translated this by postulating that concrete terms (e.g., action verbs) generically direct attention to details of an event or object. Contrastively, abstract terms (e.g., adjectives) were postulated to draw attention to the global properties of an object. In four experiments with an unrelated task paradigm we exposed participants to either abstract predicates (adjectives) or concrete predicates (action verbs). This was achieved by asking participants to use either verbs or adjectives in a spontaneous narrative or by subtly priming them with these categories supraliminally (Experiments 2 and 3) or subliminally (Experiment 4). This was followed by an ostensibly unrelated second task. In this second task, the participants were instructed to complete one or more dependent measures. These were designed to examine the attention-driving implications of the concrete versus abstract linguistic categories and consisted of either self-report (Experiments 1 and 4), Kimchi and Palmer's (1982) perceptual global-specific focus task (Experiments 1 and 4), Isen and Daubman's (1984) categorical inclusiveness task (Experiment 2), or Kitayama et al.'s (2003) framed line test (Experiment 3). Consistently across all four experiments and all dependent variables, we were able to show that abstract predicates induced a global perceptual focus, while concrete predicates induced a local perceptual focus.

This final research example diverges from the traditional approach comparing different linguistic communities to the linguistic relativity debate. Traditionally, in this debate, the focus has been specific semantic domains, such as color (e.g., Özgen, 2004), space (e.g., Majid et al., 2004), gender (e.g., Stahlberg et al., 2001) to name a few. In contrast, our [p. 323 ↓] research (Stapel and Semin, 2007) opens an entirely different way of looking at the language-cognition interface that is made possible by taking a functional

approach to language. The implications of this “new look” are wide ranging. One such implication is for the scaffold this functional view provides for the *culture-language-perception* interface. The above-summarized findings show that different linguistic categories in general (i.e., action verbs and adjectives) drive attention to different aspects of a stimulus. However, if two different linguistic communities (say Italian and Japanese) were to have differential accessibility in their daily use of adjectives and action verbs, then what would the implications of the findings reported by Stapel and Semin (2007) be? Let me illustrate with only one example.

There is recent evidence that Japanese use preferentially concrete language (i.e., verbs) rather than abstract language (i.e., adjectives) as has been shown by Maass et al. (2006). Based on these findings and those obtained by Stapel and Semin (2007), one would predict that participants from interdependent cultures with a preference for concrete language use (e.g., Japan) should make more errors in a perceptual task, such as the framed line task (Kitayama et al., 2003) compared with participants from a culture that is independent and use more abstract language habitually (e.g., Americans). This is precisely what Kitayama et al. (2003) have demonstrated. The metasemantic attention-driving function of the language approach would therefore appear to open ways of looking and linking research findings not solely within the same linguistic community, but also furnishing a way of actually grounding cross-linguistic research as well as anchoring these findings in a more general level of explanation. In concluding, I should note that these are the early steps emerging from modifying a classic question from “What is language?” to “What is language for?” and this minor modification – in my view – opens a major reorientation full of research promise.

Acknowledgments

Supported by the Royal Netherlands Academy of Arts and Sciences ISK/4583/PAH. I would like to express my thanks to Tory Higgins whose detailed and helpful comments improved the chapter considerably.

References

- Abelson, R.P. and Kanouse, D.E. (1966) Subjective acceptance of verbal generalizations . In S. Feldman (ed.) *Cognitive Consistency: Motivational Antecedents and Behavioral Consequents* , pp. 171–197. New York: Academic Press.
- Asch, S.E. (1946) Forming impressions of personality . *Journal of Abnormal and Social Psychology* , 41, 258–290. <http://dx.doi.org/10.1037/h0055756>
- Au, T.K. (1986) A verb is worth a thousand words: The causes and consequences of interpersonal events implicit in language . *Journal of Memory and Language* , 25, 104–122. <http://dx.doi.org/10.1016/0749-596X%2886%2990024-0>
- Bendix, E.H. (1966) Componential Analysis of General Vocabulary: The Semantic Structure of a Set of Verbs in English, Hindu and Japanese . The Hague: Mouton.
- Berlin, B. and Kay, P. (1969) Basic Color Terms: Their Universality and Evolution . Berkeley: University of California Press.
- Brown, R. and Fish, D. (1983) The psychological causality implicit in language . *Cognition* , 14, 233–274. <http://dx.doi.org/10.1016/0010-0277%2883%2990006-9>
- Brown, R. and Lenneberg, E.H. (1954) A study in language and cognition . *Journal of Abnormal and Social Psychology* , 49, 454–462. <http://dx.doi.org/10.1037/h0057814>
- Caramazza, A., Grober, E., Garvey, C. and Yates, J. (1977) Comprehension of anaphoric pronouns . *Journal of Verbal Learning and Verbal Behavior* , 16, 601–609. <http://dx.doi.org/10.1016/S0022-5371%2877%2980022-4>
- de Poot, C. and Semin, G.R. (1995) Pick your verbs with care when you formulate a question! . *Journal of Language and Social Psychology* , 74, 351–369. <http://dx.doi.org/10.1177/0261927X950144002>
- Douglas, K.M. and Sutton, R.M. (2006) When what you say about others says something about you: Language abstraction and inferences about described

attitudes and goals . *Journal of Experimental Social Psychology* , 42, 500–508. <http://dx.doi.org/10.1016/j.jesp.2005.06.001>

Ehrlich, K. (1980) Comprehension of pronouns . *Quarterly Journal of Experimental Psychology* , 32, 247–255. <http://dx.doi.org/10.1080/14640748008401161>

Fiedler, K. and Semin, G.R. (1989) On the causal information conveyed by different interpersonal verbs: The role of implicit sentence context . *Social Cognition* , 6, 21–39. <http://dx.doi.org/10.1521/soco.1988.6.1.21>

Fillmore, C.J. (1968) The case for case . In E. Bach, ed. and R.G. Harms (eds), *Universals in Linguistic Theory* . pp. 1–87. New York: Holt, Rinehard and Winston.

Franco, F.M. and Maass, A. (1996) Implicit versus explicit strategies of out-group discrimination: The role of intentional control in biased language use and reward allocation . *Journal of Language and Social Psychology* , 15, 335–359. <http://dx.doi.org/10.1177/0261927X960153007>

Franco, F.M. and Maass, A. (1999) Intentional control over prejudice: When the choice of the measure matters . *European Journal of Social Psychology* , 29, 469–477. <http://dx.doi.org/10.1002/%28SICI%291099-0992%28199906%2929:4%3C469::AID-EJSP938%3E3.0.CO;2-S>

Garnham, A. and Oakhill, J. (1985) On-line resolution of anaphoric pronouns: Effects of inference making and verb semantics . *British Journal of Psychology* , 76, 385–393. <http://dx.doi.org/10.1111/j.2044-8295.1985.tb01961.x>

Garnham, A., Oakhill, J. and Cruttenden, H. (1992) The role of implicit causality and gender cue in the interpretation of pronouns . *Language and Cognitive Processes* , 7, 231–255. <http://dx.doi.org/10.1080/01690969208409386>

Garnham, A., Traxler, M., Oakhill, J. and. Gemsbacher, M.A. (1996). The locus of implicit causality effects in comprehension . *Journal of Memory and Language* , 35, 517–543. <http://dx.doi.org/10.1006/jmla.1996.0028>

Garvey, C. and Caramazza, A. (1974) Implicit causality in verbs . *Linguistic Inquiry* , 5, 459–464.

Garvey, C, Caramazza, A. and Yates, J. (1976) Factors influencing assignment of pronoun antecedents . *Cognition* . 3, 227–243. <http://dx.doi.org/10.1016/0010-0277%2874%2990010-9>

Greene, S. and McKoon, G. (1994) Initiating/reacting verbs and the accessibility of discourse referents . Unpublished manuscript, Princeton University.

Gilson, C. and Abelson, R.P. (1965) The subjective use of inductive evidence . *Journal of Personality and Social Psychology* , 2, 301–310. <http://dx.doi.org/10.1037/h0022268>

Grice, H.P. (1975) Logic and conversation . In P. Cole, ed. and J. Morgan (eds), *Syntax and Semantics* , pp. 41–58. New York: Academic Press.

Hardin, C. and Banaji, M.R. (1993) The influence of language on thought . *Social Cognition* , 11, 277–308. <http://dx.doi.org/10.1521/soco.1993.11.3.277>

Hoffman, C. and Tschir, M.A. (1990). Interpersonal verbs and dispositional adjectives: The psychology of causality embodied in language . *Journal of Personality and Social Psychology* , 58, 765–778. <http://dx.doi.org/10.1037/0022-3514.58.5.765>

Hunt, E. and Agnoli, F. (1991) The Whorfian hypothesis: A cognitive psychology perspective . *Psychological Review* , 98, 377–389. <http://dx.doi.org/10.1037/0033-295X.98.3.377>

Hutchins, E. (1996) *Cognition in the Wild* . Cambridge, MA: MIT Press.

Isen, A.M. and Daubman, K.A. (1984). The influence of affect on categorization . *Journal of Personality and Social Psychology* , 47, 1206–1217. <http://dx.doi.org/10.1037/0022-3514.47.6.1206>

Kanouse, D.E. (1972) Verbs as implicit quantifiers . *Journal of Verbal Learning and Verbal Behavior* , 11, 141–147. <http://dx.doi.org/10.1016/S0022-5371%2872%2980070-7>

Kashima, Y. and Kashima, E.S. (2003) Individualism, GNP, climate, and pronoun drop: Is individualism determined by affluence and climate, or does language use play a role? *Journal of Cross Cultural Psychology* , 34, 125–134. <http://dx.doi.org/10.1177/0022022102239159>

Kasoff, J. and Lee, J.Y. (1993) Implicit causality as implicit salience . *Journal of Personality and Social Psychology* , 65, 877–891. <http://dx.doi.org/10.1037/0022-3514.65.5.877>

Kay, P. (1996) Inter-speaker relativity . In J.J. Gumpertz, ed. and S.C. Levinson (eds), *Rethinking Linguistic Relativity* pp. 97–114. Cambridge: Cambridge University Press.

Kimchi, R. and Palmer, S.E. (1982) Form and texture in hierarchically constructed patterns . *Journal of Experimental Psychology: Human Perception and Performance* , 8, 521–535. <http://dx.doi.org/10.1037/0096-1523.8.4.521>

Kitayama, S., Duffy, S., Kawamura, T. and Larsen, J.T. (2003) Perceiving an object and its context in different cultures . *Psychological Science* , 14, 201–207. <http://dx.doi.org/10.1111/1467-9280.02432>

Lenneberg, E.H. (1953) Cognition in ethnolinguistics . *Language* , 29, 463–471. <http://dx.doi.org/10.2307/409956>

Lenneberg, E.H. and Roberts, J.M. (1956) The language of experience: A study in methodology . *International Journal of American Linguistics* , 22 (2, part 2, Memoir 13).

Lippman, W. (1922) *Public Opinion* . New York: Hartcourt, Brace and Co.

Maass, A. and Arcuri, L. (1992) The role of language in the persistence of stereotypes . In G. Semin, ed. and K. Fiedler (eds), *Language, Interaction and Social Cognition* , pp. 129–43. Newbury Park, CA: Sage.

Maass, A., Karasawa, M., Politi, F. and Suga, S. (2006) Do verbs and adjectives play different roles in different cultures? A cross-linguistic analysis of person representation . *Journal of Personality and Social Psychology* , 90, 734–750. <http://dx.doi.org/10.1037/0022-3514.90.5.734>

Maass, A., Milesi, A., Zabbini, S. and Stahlberg, D. (1995) The linguistic intergroup bias: Differential expectancies or in-group protection? *Journal of Personality and Social Psychology* , 68, 116–126. <http://dx.doi.org/10.1037/0022-3514.68.1.116>

Maass, A., Salvi, D., Arcuri, L. and Semin, G. (1989) Language use in intergroup contexts: The linguistic intergroup bias . *Journal of Personality and Social Psychology* , 57, 981–993. <http://dx.doi.org/10.1037/0022-3514.57.6.981>

Majid, A., Bowerman, M., Kita, S., Haun, D.B.M. and Levinson, S.C. (2004) Can language restructure cognition? The case for space . *Trends in Cognitive Sciences* , 8, 108–114. <http://dx.doi.org/10.1016/j.tics.2004.01.003>

McKoon, G., Geene, S.B. and Ratcliff, R. (1993) Discourse models, pronoun resolution, and the implicit causality of verbs . *Journal of Experimental Psychology: Learning, Memory and Cognition* , 19, 1040–1052. <http://dx.doi.org/10.1037/0278-7393.19.5.1040>

Miller, G.A. and Johnson-Laird, P. (1976) Language and Perception . Cambridge: Cambridge University Press.

Özgen, E. (2004) Language, learning, and color perception . *Current Directions in Psychological Science* , 13, 95–98.

De Poot, C. and Semin, G.R. (1995) Pick your verbs with care when you formulate a question! *Journal of Language and Social Psychology* , 14, 351–368. <http://dx.doi.org/10.1177/0261927X950144002>

Reitsma-van Rooijen, M., Semin, G.R. and van Leeuwen, E. (2007) The effects of linguistic abstraction on interpersonal distance . *European Journal of Social Psychology* , 37, 817–823. <http://dx.doi.org/10.1002/ejsp.410>

Rubini, M. and Kruglanski, A.W. (1997) Brief encounters ending in estrangement: Motivated language use and interpersonal rapport in the question-answer paradigm . *Journal of Personality and Social Psychology* , 72, 1047–1060. <http://dx.doi.org/10.1037/0022-3514.72.5.1047>

Rudolph, U. and Försterling, F. (1997) The psychological causality implicit in verbs: A review . *Psychological Bulletin* , 121, 192–218. <http://dx.doi.org/10.1037/0033-2909.121.2.192>

Sapir, E. (1951) In D.G. Mandelbaum (ed.), *Selected Writings of Edward Sapir in Language, Culture, and Personality* . Berkeley: California University Press.

Semin, G.R. (1989) The contribution of linguistic factors to attribute inferences and semantic similarity judgments . *European Journal of Social Psychology* , 19, 85–100. <http://dx.doi.org/10.1002/ejsp.2420190202>

Semin, G.R. (1995) Interfacing language, and social cognition . *Journal of Language and Social Psychology* , 14, 182–194. <http://dx.doi.org/10.1177/0261927X95141010>

Semin, G.R. (1998) Cognition, language, and communication . In S.R. Fussell, ed. and R.J. Kreuz (eds), *Social and Cognitive Psychological Approaches to Interpersonal Communication* , pp. 229–257. Hillsdale, NJ: Laurence Erlbaum.

Semin, G.R. (2000a) Agenda 2000: Communication: language as an implementational device for cognition . *European Journal of Social Psychology* , 30, 595–612. <http://dx.doi.org/10.1002/1099-0992%28200009/10%2930:5%3C595::AID-EJSP23%3E3.0.CO;2-A>

Semin, G.R. (2000b) Language as a cognitive and behavioral structuring resource: Question-answer exchanges . In W. Stroebe, ed. and M. Hewstone (eds), *European Review of Social Psychology* , pp. 75–104. Chichester: Wiley.

Semin, G.R. (2001) Language and social cognition . In A. Tesser, ed. and N., Schwarz (eds), *Handbook of Social Psychology, Intraindividual Processes* , 1, 159–180. Oxford: Blackwell.

Semin, G.R. (2006) Modeling the architecture of linguistic behavior: Linguistic compositionality, automaticity, and control . *Psychological Inquiry* , 17, 246–255.

Semin, G.R. (2007) Grounding communication: Synchrony . In A. Kruglanski, ed. and E.T. Higgins (eds), Social Psychology: Handbook of Basic Principles , 2nd Edition, pp. 630–649. New York: Guilford Publications.

Semin, G.R. and Chassein, J. (1985) The relationship between higher order models and everyday conceptions of personality . European Journal of Social Psychology , 15, 1–15. <http://dx.doi.org/10.1002/ejsp.2420150102>

Semin, G.R. and Fiedler, K. (1988) The cognitive functions of linguistic categories in describing persons: Social cognition and language . Journal of Personality and Social Psychology , 54, 558–568. <http://dx.doi.org/10.1037/0022-3514.54.4.558>

Semin, G.R. and Fiedler, K. (1989) Relocating attributional phenomena within a language-cognition interface: The case of actors' and observers' perspectives . European Journal of Social Psychology , 19, 491–508. <http://dx.doi.org/10.1002/ejsp.2420190602>

Semin, G.R. and Fiedler, K. (1991) The linguistic category model, its bases, applications and range . In W. Stroebe, ed. and M. Hewstone (eds), European Review of Social Psychology , 2, 1–30. Chichester: Wiley. <http://dx.doi.org/10.1080/14792779143000006>

Semin, G.R. and Fiedler, K. (1992) The configuration of social interaction in interpersonal terms . In G.R. Semin, ed. and K. Fiedler (eds), Language, Interaction and Social Cognition . London: Sage Publications.

Semin, G.R., Gorts, C, Nandram, S. and Semin-Goossens, A. (2002) Cultural perspectives on the linguistic representation of emotion and emotion events . Cognition and Emotion , 16, 11–28. <http://dx.doi.org/10.1080/02699930143000112>

Semin, G.R. and Greenslade, L. (1985) Differential contributions of linguistic factors to memory based ratings: Systematizing the systematic distortion hypothesis . Journal of Personality and Social Psychology , 49, 1713–1723. <http://dx.doi.org/10.1037/0022-3514.49.6.1713>

Semin, G.R. and Krahé, B. (1987) Lay conceptions of personality: Eliciting tiers of a scientific conception of personality . European Journal of Social Psychology , 17, 199–209. <http://dx.doi.org/10.1002/ejsp.2420170207>

Semin, G.R. and Krahé, B. (1988) Explaining perceived cross-situational consistency: Intuitive psychometrics or semantic mediation? European Journal of Personality , 2, 239–252. <http://dx.doi.org/10.1002/per.2410020402>

Semin, G.R. and De Poot, C.J. (1997a) The question-answer paradigm: You might regret not noticing how a question is worded . Journal of Personality and Social Psychology , 73, 472–480. <http://dx.doi.org/10.1037/0022-3514.73.3.472>

Semin, G.R. and De Poot, C.J. (1997b) Bringing partiality to light: Question wording and choice as indicators of bias . Social Cognition , 15, 91–106. <http://dx.doi.org/10.1521/soco.1997.15.2.91>

Semin, G.R. and Smith, E.R. (2008) Embodied Grounding: Social, Cognitive, Affective, and Neuroscientific Approaches . New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511805837>

Semin, G.R., Rubini, M. and Fiedler, K. (1995) The answer is in the question: The effect of verb causality on locus of explanation . Personality and Social Psychology Bulletin , 21, 834–841. <http://dx.doi.org/10.1177/0146167295218006>

Smith, E.R. and Semin, G.R. (2004) Socially situated cognition: Cognition in its social context . Advances in Experimental Social Psychology , 36, 53–117. <http://dx.doi.org/10.1016/S0065-2601%2804%2936002-8>

Stahlberg, D., Sczesny, S. and Braun, F. (2001) Name your favorite musician: Effects of masculine generics and of their alternatives in German . Journal of Language and Social Psychology , 20, 464–469. <http://dx.doi.org/10.1177/0261927X01020004004>

Stapel, D. and Semin, G.R. (2007) The magic spell of language . Journal of Personality and Social Psychology (in press).

Von Hippel, W., Sekaquaptewa, D. and Vargas, P. (1997) The linguistic intergroup bias as an implicit indicator of prejudice . *Journal of Experimental Social Psychology* , 33, 490–509. <http://dx.doi.org/10.1006/jesp.1997.1332>

Von Hippel, W., Sekaquaptewa, D. and Vargas, P. (2007) Armed only with paper and pencil: 'low-tech' measures of implicit attitudes . In B. Wittenbrink, ed. and N. Schwarz (eds), *Implicit Measures of Attitudes: Procedures and Controversies* , pp. 103–124. New York: Guilford Press.

von Humboldt, W. (1836) Über die Verschiedenheit des menschlichen Sprachaufbaues und ihren Einflu# auf die geistige Entwicklung des Menschengeschlechts . Berlin: Royal Academy of Sciences: Republished (trans. G.C. Buck, ed. and F. Raven, ed.) (1999) *Linguistic Variability and Intellectual Development* . Baltimore, MD: University of Miami Press.

Werkman, W.M., Wigboldus, D.H. and Semin, G.R. (1999) Children's communication of the Linguistic Intergroup Bias and its impact upon cognitive inferences . *European Journal of Social Psychology* , 29, 95–104. <http://dx.doi.org/10.1002/%28SICI%291099-0992%28199902%2929:1%3C95::AID-EJSP898%3E3.0.CO;2-Z>

Whorf, B.L. (1956) *Language, Thought, and Reality* Cambridge, MA: MIT Press.

Wigboldus, D., Semin, G.R. and Spears, R. (2000) How do we communicate stereotypes? Linguistic bases and inferential consequences . *Journal of Personality and Social Psychology* , 78, 5–18. <http://dx.doi.org/10.1037/0022-3514.78.1.5>

<http://dx.doi.org/10.4135/9781446249215.n16>

Handbook of Theories of Social Psychology: Volume 1

Action Identification Theory

Contributors: Robin R. Vallacher & Daniel M. Wegner

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Action Identification Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n17>

Print pages: 327-349

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n17>

[p. 327 ↓]

Chapter 16: Action Identification Theory

Robin R. Vallacher, ed. and Daniel M. Wegner, ed.

Abstract

The theory of action identification is a system of three principles explaining how people's thoughts of what they are doing relate to what they do. In a sense, the principles suggest an "operating system" for a human being – a program linking thought to action. The principles suggest simply that people do what they think they are doing, that when they can think of doing something more, they do that – but that when they can't do what they were thinking of doing, they think of doing less. Stated this way, the principles may seem perhaps too simple, but their interaction and ramifications are remarkably complex and predictive of a rich array of phenomena. The theory has things to say about how people can conceptualize their actions optimally, it offers insights into how people regulate actions through thinking, it provides a template for understanding how action connects to emotion and to self-concept, and it suggests how social influence occurs by changing how actions are understood. These consequences of the theory were discovered early on, but have recently been supplemented by extensions to encompass the dynamics of action and the role of action identification in the understanding of own and others' minds.

Introduction

Life is full of big and little things. One moment you might be planning your career; the next moment you might be turning on your computer or searching for a pen. Sometimes

the little things are fairly automatic, requiring little conscious attention, and are irrelevant to the bigger things we have in mind. And sometimes the little things are simply subordinate components to the big things, the details by which we accomplish goals, implement plans, or demonstrate traits and values. So although everyday life can be characterized as the ebb and flow of actions, big and little, psychologists typically ignore the little ones and concentrate instead on people's broader action tendencies. Theories in social psychology are replete with references to goals, values, personal standards, schemas, chronic concerns, and personality traits. Searching for a pen or turning on a computer may be unavoidable details of everyday life but not all that significant in their own right.

[p. 328 ↓] The theory of action identification (Vallacher and Wegner, 1985, 1987; Wegner and Vallacher, 1986) grants a far bigger role to the little things. Or more precisely, the theory posits a deep and reciprocal connection between the details of what we do and the larger meanings we have in mind. This dynamic interplay between mind and action is important in its own right, but it is also highly germane to a number of issues of concern to social psychologists. Not all of these implications were on the radar screen for us when we developed the theory in the 1980s. Our aim at the time was far more circumscribed: to generate a parsimonious model of the mental control of action. Since that time, the theory has evolved in two ways. First, it has been extended to topics that go beyond mental control per se, from personality and self-concept to social influence and conflict resolution. Second, the theory has morphed into forms that resonate with contemporary perspectives in psychology, including cognitive science and dynamical systems. The emergent nature of the theory is apparent in some of the work the two of us have pursued independently, often without explicit reference to the original theory.

Our big aim in this chapter is to describe action identification theory and trace its trajectory from the 1980s to the present. We begin by outlining the basic theory and its initial range of application. We then indicate how the theory's scope has broadened over the years to account for important domains of social experience. In the next part, we discuss how the theory has been reframed in recent years in light of new developments in psychology and science generally. In a concluding section, we throw caution to the wind and suggest that the dynamic scenario at the heart of the theory provides a template for human experience. When liberated a bit from its specific principles

and couched in more general terms, action identification theory can be viewed as a metatheory that expresses the essence of personal and interpersonal dynamics.

The Basic Theory

Back in the early and mid 1980s, we were blissfully unaware of where our actions would take us. We were certainly trying to think big, but we were also caught up in lots of little things. In big terms, we looked upon our collaborative effort as constructing a theory about mental control. We also saw our mission as integrating insights from several perspectives that were coming into their own at the time. Attribution theory and cognitive dissonance phenomena were never far from our minds, but of more immediate interest were cybernetic models of action (Carver and Scheier, 1981), unitization in behavior perception (Newton, 1976), and the endogenous-exogenous distinction in lay epistemology (Kruglanski, 1975). Our reading and thinking even strayed from psychology to encompass philosophical approaches to the understanding of action – like those of Goldman (1970) and Hampshire (1960). Each of these perspectives conveyed something fundamental about the relationship between mind and action, and we aspired to synthesize these insights in a parsimonious framework.

But we also did lots of little things, many of which seemed at odds with constructing a new theory. We stayed up late, occasionally drank scotch, made goofy jokes, offered wild-eyed speculations on the human condition, and even constructed a few faux psychological instruments (e.g., the Hidden Brain Damage Scale) to make each other (and others with suspicious standards) laugh. Out of these seemingly irrelevant asides, we experienced a number of insights and breakthroughs that transformed our thinking about mind and action – although these probably paled in number with the all-too-frequent “What were we thinking!?” false starts. In a reflexive approach to theory construction that borders on narcissism, our reflections on the role of these little acts in promoting a bigger sense of what we were doing became central to the theory itself. We realized that an action, even a highly important one like [p. 329 ↓] theory construction, tends to become transformed mentally as well as mechanically over time and we suspected that there might be a reliable and meaningful pattern to such transformations. Perhaps, we thought, the movement between big and little things is central to the mind-action relationship, satisfying the dual and sometimes conflicting criteria of effectiveness

and global understanding. At that point, of course, we promptly opened another bottle of scotch.

Central to our thinking at the time was the inherently ambiguous nature of human action and the perils this posed for theory construction. But we came to see this issue as a promise, more than a dilemma, and proceeded to develop the theory around it. Below we describe this progression from the uncertainty of action to the certainty of action understanding conveyed in the theory's three principles.

The Uncertain Act

In science, the phenomenon to be explained is usually well defined. When chemists theorize about chemical reactions, they don't have to guess about the nature of molecules. The same certainty lies at the heart of physics. Sure, quarks and dark energy are inferred, but at least the inferences, however controversial, are clear and precise. Even in the social sciences, the fundamental units are well defined. Economists develop models concerning the flow of money and tangible resources, anthropologists construct theories about religious beliefs and cultural traditions, and political scientists theorize about power structures and voting practices.

Social psychology is certainly a science, but the building blocks for theory construction in our field are not as well defined as are the building blocks in other scientific realms. Social psychological theories focus on a wide variety of constructs – norms, roles, beliefs, values, and so forth – but underlying all these constructs is the concept of action. Social psychology is ultimately concerned with what people *do*, whether privately or in explicit social contexts such as relationships, groups, and crowded streets. But therein lies the rub: what exactly is the unit of action that corresponds to the physicist's protons, the biologist's genes, or the economist's dollar? What is an action and how can it be defined unequivocally so that it can provide a solid foundation for theory construction?

This question, we soon learned, was a bit of an obsession in philosophy. Coming from distinct traditions and armed with diverse assumptions, many philosophers have noted that an action does not admit to a single, unequivocal definition, but rather can

be identified in many different ways (e.g., Danto, 1963; Goldman, 1970; Ryle, 1949). Something as straightforward as “driving to work” could be just as accurately identified as “operating a motor vehicle,” “burning gas,” “obeying traffic laws,” “following a daily routine,” or simply as “stepping on pedals” and “turning a wheel.” When the focus turns to actions that are of greater interest to social psychologists – harmful versus helpful behavior, fairness versus injustice, and so on – the availability of multiple identities for the same action can prove problematic for theory construction. The act of “criticizing someone,” for instance, could be viewed by observers, or even by the actor him or herself, as “acting rudely,” “asserting dominance,” “offering a different point of view,” “providing constructive feedback,” or simply as “uttering words.” Which identity is deemed the “real” one makes a big difference in the explanation that is generated.

Even if there is consensus in the meaning attached to someone's action (e.g., the critical comment is widely perceived to be mean-spirited), the action still does not have an unequivocal definition. The same action in a different context might promote a very different consensus regarding its meaning (e.g., the critical comment in a policy-making group might be seen as constructive). It is also the case that actions are commonly characterized by equifinality (functional equivalence), further complicating the task of establishing stable building blocks for [p. 330 ↓] theory construction. Mean-spirited, for example, provides functional equivalence for both verbal criticism and the silent treatment. Criticism, in turn, can be instantiated in very different ways (e.g., pointed comments or sarcasm). This problem exists even if one reduces action to button pushes or other simple operational definitions. A button can be pushed by a wide variety of physical motions involving different limbs and digits. Beyond that, when an action is reduced in this fashion, the number of potential identities for it increases in a proportional manner. A button push can mean virtually anything depending on the program to which it is attached or the context in which it occurs. Protons and dollars don't pose this problem.

To a critic, it might seem that social psychology is built on shifting sands, without a clear foundation for theory construction. Just such a criticism was in ascendance during the period in which we were busy developing our theory. Gergen (1985) was especially vocal in this regard, arguing that a science of social psychology was impossible in principle because of the inherent ambiguity of action. He, and others who shared this sentiment, seemed to suggest that we should give up the pretense and concentrate

instead on how to make people happy (whatever that is!). But we were mindful of Kurt Lewin's remark that "one man's artifact is another man's theory." In this spirit, it occurred to us that far from being a problem for theory construction in psychology, the inherent ambiguity of action could be viewed as the key to understanding the relationship between mind and action. After all, despite the multiple identities available for any action, people seem to know in unequivocal terms what they are doing, have done, or intend to do. Somehow people sidestep the uncertainty of action and navigate the stream of potential act identities one at a time. How do they do it?

Levels of Action Identification

Ironically, we can make sense of action certainty by looking into the human mind that is ultimately responsible for the ambiguity of action. The mind is designed to identify (or create) patterns in the real world. Lacking this pervasive tendency, people would be overwhelmed by the complex and nuanced information that continually bombards their senses on a moment-to-moment basis. Meaningful actions exist because we find or impose patterns on the specific behaviors we observe or otherwise learn about. The patterns are constructions, but once generated, they are maintained because they disambiguate reality and thereby provide coherent understanding and a stable platform for subsequent thought and behavior. Because they are constructions, however, they can admit to tremendous variability across people and contexts. Hence, the certainty of action that exists for each individual embedded in a particular context coexists with the uncertainty of action across individuals and contexts. Philosophers and psychologists live in the latter world, at least when wearing their professional hats and when talking about other people's realities.

That said, there is one metric for disambiguating action that seems solid and reflects a shared reality. The multiple act identities for an action tend to be organized in a hierarchical manner. Lower-level identities in the hierarchy convey the details or specifics of the action and thus indicate *how* the action is performed. Higher-level identities provide a more general understanding of the action; they indicate *why* the action is performed or what its effects and implications are. Higher-level identities are less movement-defined than lower-level identities, and provide a more abstract and comprehensive understanding of the action. Identification level is relative, so whether

a particular act identity is a high or low level depends on the identity with which it is compared. What looks like a goal (e.g., “getting married”) can be identified as a means with respect to a higher level identity (e.g., “starting a family”) An act identity that seems molecular (e.g., “turning on a computer”), meanwhile, could represent a goal with respect to a yet lower-level identity (e.g., “pushing a button”).

[p. 331 ↓] A simple criterion is useful for sorting an action's multiple identities into a hierarchy. One act identity is higher-level than another identity if it makes sense to say that one does the former *by* doing the latter (Goldman, 1970). Thus, one “goes to work” *by* “driving a car,” one “drives a car” *by* “stepping on pedals and turning the steering wheel,” and one “steps on pedals and turns the wheel” *by* “moving one's arms and leg.” People appreciate the notion of an asymmetric *by* relation and demonstrate a very high degree of consensus in sorting act identities hierarchically in this fashion (Vallacher and Wegner, 1985).

Some act identities are not asymmetric in this sense, but rather have a *level-indeterminate* relation with one another. “Being rude” and “being constructive” are both plausible higher-level identities for “criticizing someone,” for example, but they do not have an obvious hierarchical relation. To distinguish among such act identities in a meaningful manner, we have employed various psychometric methods, including multidimensional scaling and, more commonly, factor analysis. In the factor analytic approach, participants rate a wide variety of identities for an action (generated in pilot research) according to how well they personally describe the action. Factor analyses of these ratings typically reveal a single low-level factor that captures the most rudimentary act identities and two or more higher-level factors that reflect differences among identities with respect to valence and content-relevant dimensions. “Being rude,” for example, loads on the same factor as other negative implications of criticism (e.g., “showing disrespect,” “disregarding someone's point of view”), whereas “being constructive” loads on the same factor as other positive higher-level identities of criticism (e.g., “offering a different point of view,” “providing useful feedback”). The dimensionality of higher-level act identities has theoretical relevance for a host of personal and social phenomena. These implications follow from the principles of action identification.

Theoretical Principles

It is not surprising that people can distinguish among an action's available identities by virtue of an asymmetric *by relation*. But how do people pick one level over another to identify what they are doing, have done, or intend to do? Why does someone look upon his or her action as, say, "going to work" rather than "driving a car," or as "playing a piano" rather than "expressing feelings"? And after an action is identified in a particular way, what determines whether it is maintained under that identity or instead is reidentified later in different terms? Why do higher-level identities (e.g., "being cooperative") sometimes resist change in the face of social pressure, but at other times change dramatically in content and valence (e.g., "being competitive")? What is the relationship between how an action is identified and how effectively the action is performed? The answers to these questions are conveyed in three principles.

Principle 1: Action is Maintained with Respect to its Prepotent Identity

Action identification would hardly be worthwhile for people, let alone worth theorizing about by social psychologists, in the absence of this principle. It is central to models of self-regulation, of course, and in a broader sense to any theory that posits a link between mind and action (e.g., Carver and Scheier, 2002; Higgins, 1998; James, 1890; Miller et al., 1960). A person's prepotent act identity functions as an intention to initiate an action, a frame of reference for performing the action, and a criterion to assess how well the action has been performed. Because an action can be identified at different levels, the principle directly implies that people can maintain action at different levels. A person may intend to "send an e-mail," for instance, and monitor his or her subsequent action to see whether this intention was fulfilled. Alternatively, the person may intend to "contact a colleague" (a higher-level act identity) or "propose a collaborative writing venture" (a yet higher-level act identity), and monitor [p. 332 ↓] the attainment of whichever act identity is prepotent.

Principle 2: When Both a Lower- and a Higher-Level Act Identities are Available, There is a Tendency for the Higher-Level Identity to Become Prepotent

The idea here is that people are sensitive to the meanings, consequences, and implications of what they do. Within social psychology, this principle is reflected in the emphasis on goals, plans, values, and other global constructs that are said to motivate personal and interpersonal behavior. The preference for higher-level understanding is not unique to social psychology, but rather is central to many theoretical traditions, including learning under reinforcement contingencies, the mastery of skilled action, pattern formation and recognition in Gestalt psychology, inductive reasoning in cognitive psychology, and the “search for meaning” in existential psychology. These different schools of psychology share the assumption that act representations expand to encompass broader effects and meanings. In learning, a basic act expands to incorporate the reinforcing effects of the action. In the development of mastery, discrete acts become automated and integrated into a larger action unit. In Gestalt psychology, parts are unified to produce a perceptual whole. In induction, separate observations or pieces of information are organized into an explanatory account. And in existentialism, patterns discerned in distinct actions are recognized as manifestations of a larger action tendency.

Principle 2 enables people to choose one of many plausible act identities and actually do something. Without this principle, people would be “buried in thought,” entertaining a multitude of possible intentions in a given context. So when two or more plausible identities are available, people are inclined to choose the identity that provides the most comprehensive understanding of what they are doing, plan to do, or have done. A person could think about his or her behavior as “using eating utensils,” for example, but is more inclined to gloss over such details and identify the behavior as “eating dinner.” If cues to yet higher-level identities are available, the person is likely to embrace them

over the now-lower-level “eating identity.” Thus, he or she may look upon the meal as “getting nutrition,” “satisfying my appetite,” or “putting on weight.”

Principle 3: When an Action Cannot Be Performed in Terms of its Prepotent Identity, There is a Tendency for a Lower-Level Identity to Become Prepotent

If Principle 2 was the only basis for action identification, people's minds would be populated by abstractions, fantasies, hopes, and fears, as increasingly higher-level identities emerge as ways of thinking about one's action. Even the most rudimentary act could be charged with high-level significance in this manner. Such progressive integration would be possible if people lived in a world in which every thought was feasible and easily enacted.

Reality is not so accommodating. There are obstacles to enacting goals and plans, and even in the absence of such disruptions, the personal difficulty of achieving one's ends can derail an action undertaken with respect to a higher-level identity. A person might set out to “demonstrate tennis prowess,” for example, only to disappoint him or herself, as well as observers, as he or she loses point after point during the tennis match. To regain control of the action, the person is inclined to adopt a more manageable lower-level act identity, such as “hitting cross-court shots.” If this act identity is not effectively performed, the person might drop to a yet Lower-level act identity, thinking about “preparing the racket” or “getting in proper position before swinging the racket.” Whereas Principle 2 can lead to progressively higher levels of action identification, Principle 3 pulls people in the opposite direction, leading to progressively finer gradations of detail, with attention devoted to increasingly molecular features of what they are doing. The potential for flights of fancy that is inherent in the second principle, then, is unlikely to [p. 333 ↓] pose a serious problem for most people because of the reality orientation inherent in the third principle.

Several factors dictate how detailed a person's act identity is likely to be. Actions that are complex, unfamiliar, or time-consuming tend to be identified in low-level terms. An easy or familiar action may also be identified in lower-level terms if it is disrupted. And any action – even one that is fairly easy and free from disrupting influences – tends to be identified in lower-level terms as the time for enactment approaches, a point that has been further documented in temporal construal theory (Trope and Lieberman, 2003). Engaged couples, for example, identify the act of getting married in high-level terms several weeks in advance of the wedding, but think about progressively finer details (e.g., saying vows, walking down the aisle) as the day approaches (Vallacher and Wegner, 1985: Chapter 4). Finally, simply instructing people to think about their action in molecular terms can prove sufficient to induce a low-level mindset.

The Emergence Process

Taken together, the three principles impart a dynamic interplay to the connection between mind and action. Low-level identification is a relatively unstable state that is adopted out of necessity rather than preference. Movement to a lower-level state (Principle 3) thus provides the precondition for adopting a higher-level identity (Principle 2) that restores coherence in the mental system and provides a stable platform for action (Principle 1).

Sometimes this process is akin to getting around a roadblock. After a brief detour to lower-level details, the person is back on track implementing the original higher-level identity. If this were always the case, though, people would never develop new ways of acting. But people do develop new insights into their actions and often chart new courses of action. When a higher-level meaning has been abandoned in order to regain control of an action at a lower level, for example, the person becomes sensitive to cues to higher-level meaning in the action context, and these may provide an avenue of emergence to a new way of understanding the action. Lacking the lower-level state, the change from one high-level identity to a different one would not occur.

Our initial experiments on action identification were attempts to validate the emergence process. Wegner et al. (1984, Experiment 1), for example, investigated whether experienced coffee drinkers could be led to think about coffee drinking in a new way if

they were induced to focus on the details of drinking. Some participants read an essay arguing that coffee drinking makes people *seek out stimulation*; others read an essay arguing instead that coffee drinking makes people *avoid stimulation*. After reading the essay, participants listened to musical passages through headphones, with the expectation that they would subsequently rate the passages. They were free to adjust the volume of the music if it was too loud or not loud enough. Half the participants drank from normal coffee cups, but the others drank from heavy cups covered with duct tape (ostensibly as a safeguard against electric shock because of the electronic equipment in the music appreciation task).

Participants who drank coffee from the normal cups did not adjust the volume reliably one way or the other in response to the essay they read. Presumably, they already had a high-level identity for coffee drinking (which we verified in a pilot study), so the identity for coffee drinking provided in the essay did not provide a new way to think about the act. But participants who drank coffee from the unwieldy cups were influenced by the essay: those who read the “seeking stimulation” essay turned the volume up (increasing stimulation), whereas those who read the “avoiding stimulation” essay turned the volume down (decreasing stimulation). The unwieldy cups induced lower-level act identities for coffee drinking (verified in pilot research), such as “lifting a [p. 334 ↓] cup” and “swallowing liquid.” The essay then provided an avenue of emergence to a higher-level identity (seeking versus avoiding stimulation) of what they were doing.

Emergence also takes place on a longer timescale. As people become increasingly competent at an action, for example, they tend to identify the action in terms of its consequences, self-evaluative implications, and other forms of meaning, rather than in terms of its lower-level details. This “sealing off” of lower-level act identities, which is consistent with research on skill acquisition, has been demonstrated for a variety of actions, including piano playing, essay writing, tennis, karate, and videogames (Vallacher and Wegner, 1985). In all cases, people initiate the act with a relatively high-level identity in mind, move to lower-level identities as they learn the action, and then move to a higher-level act identity as the action becomes more-or-less automated and mastered. We discovered, however, that the emergent identity was rarely the same high-level identity that motivated the people to begin with. Playing the piano, for example, was identified initially by many people as “impressing my friends” but after a sustained period of low-level maintenance, the now-proficient piano players came to

identify piano playing as “relaxing myself.” The tendency for an emergent act identity to differ from the action’s antecedent identity suggests a scenario by which people develop new motives, interests, concerns, and insights into their mental make-up.

Negative high-level identities can also be embraced in this fashion. People can deflect an undesirable characterization of their behavior through a variety of interchangeable cognitive mechanisms (cf. Tesser et al., 1996) as long as they have a more flattering depiction available at the same identification level. Someone informed that he or she has acted rudely or demonstrated insensitivity, for example, may be uninfluenced by this feedback if he or she looks upon the action in question as offering constructive feedback. But if the person is induced to think about the lower-level aspects of the action, he or she is primed for emergence and thus is more likely to accept the unflattering higher-level characterization. In this way, people are capable of accepting responsibility for actions with negative consequences and implications, and are open to new insights into their motives and personality dispositions.

The Optimality Hypothesis

The tension between preference (Principle 2) and necessity (Principle 3) is manifest as an oscillation over time between higher and lower levels of identification. This oscillation eventually dissipates, with the prepotent act identity converging on a restricted range of identities at a level that provides a balance between the two tendencies. This level represents the *optimal level of identification*. A person may prefer to think about his or her behavior at a party as “demonstrating wit and charm,” for example, but reality constraints may promote a lower-level orientation involving such identities as “think of funny comments,” “smile and maintain eye contact,” and “look for gullible people.” The optimal level represents a compromise between comprehensive understanding and effective action. As such, it signifies the base-rate difficulty of the action for people generally, or the level of action mastery for an individual.

Despite the tendency toward optimality in action identification, people's mental dynamics are not always in perfect resonance with their overt behavior. Indeed, the lay (and scientific) fascination with psychology is attributable in large part to problems in the feedback between mind and action conveyed in the optimality hypothesis.

People routinely fail to do what they intend, and often fail to profit from their mistakes in subsequent planning and behavior. To a certain extent, lapses in mental control reflect a lack of skill or experience with respect to the action in question. If failures in action signified only a lack of preparation or ability, however, the interest in psychology would center [p. 335 ↓] primarily on issues of learning and skill development, with the study of personality and social psychology occupying a subsidiary role. Fortunately for our field, people make mistakes despite having the requisite skill, experience, and motivation to perform the action effectively (cf. Baumeister and Heatherton, 1996).

Dysfunctional action occurs when people monitor and control what they are doing with respect to a level of identification that is non-optimal for the action's performance. This potential exists because an action's available identities are constrained by the context in which the action occurs. Most of the contexts defining everyday life are stacked in favor of relatively high-level identities, since there are usually salient cues to an action's causal effects, socially labeled meanings, and potential for self-evaluation. When a person is offered a reward or threatened with punishment, for instance, it may prove impossible for him or her not to define what he or she is doing in these terms. In similar fashion, situations involving competition, audience evaluation, or other pressures to do well may keep the person mindful of high-level identities of a self-evaluative nature (e.g., "demonstrating my skill," "trying to win," "impressing others") at the expense of the action's more molecular representations. If the action is personally difficult, the context-induced high-level identities will lack sufficient detail and coordination of components necessary for optimal performance.

This reasoning is consistent with research demonstrating that unfamiliar or complex actions are adversely affected by factors that charge the action with significance. Thus, a complex task suffers when it is linked to salient rewards, when it is performed in the presence of an evaluative audience, and when it is performed in a competitive context (cf. Baumeister and Heatherton, 1996). These factors each impart a fairly high-level representation to the task (earning a reward, impressing an audience, demonstrating one's skill), so it is not surprising that they impair performance when the task is difficult or unfamiliar and hence best enacted with respect to more molecular representations. In effect, the basic action elements are discharged without the degree of conscious control necessary to assure their moment-to-moment coordination.

On the other hand, some contexts can induce a level of action control that is too *low* for effective performance. This is likely to occur when there are distractions, obstacles, or other sources of disruption that render the action's details prepotent. A jammed keyboard, for instance, can reduce "preparing a chapter" to a series of discrete keystrokes. Low-level identities also become prepotent in novel settings lacking familiar cues to an action's higher-level meaning. In yet other contexts, a person's attention may be drawn to the details of his or her behavior, inducing him or her to experience a lower level of identification than would normally be the case. If the action is personally easy for the person and thus best performed with relatively high-level identities in mind, the prepotence of lower-level identities resulting from disruption, novelty, or instruction can undermine the quality of the person's performance.

Consistent with this possibility, research has shown that performance can be disrupted when attention is drawn to the over-learned details of an action (e.g., Langer and Imber, 1979). The more over-learned the action is, the greater the performance impairment engendered by a conscious concern with how to perform the act (Kimble and Perlmutter, 1970). A proficient piano player, for instance, can become derailed in the middle of a familiar piece if he or she looks at the music and starts to think about which sequence of keys to hit with which fingers. This form of non-optimality represents consciousness trying to micromanage a problem that is best left to lower-level echelons to work out among themselves.

Support for both manifestations of nonoptimality is provided in a study of speech fluency (Vallacher et al., 1989). Participants were asked to deliver a prepared speech to [p. 336 ↓] either an easy-to-persuade audience or a difficult-to-persuade audience. Half the participants were induced to think about the action in high-level terms ("try to persuade the audience"), and half were induced to focus on the lower-level details of delivering the speech (monitoring their voice quality). In line with the optimality hypothesis, the participants made fewer speech errors when their level of identification matched the personal difficulty of the act. Specifically, speech fluency was maximized when the task was personally easy (i.e., the audience was believed to be easy to persuade) and identified at high level and when the task was personally difficult (the audience was hard to persuade) and identified at low level. When the match between the action's difficulty and level of identification was nonoptimal (the easy act identified in low-level terms and the difficult act identified in high-level terms), participants tended to stumble over their

words and experience frequent speech pauses. The nonoptimal participants also rated their speech performance as relatively poor compared with participants in the optimality conditions.

The Extended Theory

The theory in its original form was intended to address the ambiguous nature of action and to develop principles concerning the mental control of action. It soon became apparent, however, that the reciprocal feedback between mind and action had implications for a variety of other topics of interest to social psychologists, including emotion, personality, self-concept, and social influence.

Emotion

Perhaps the most widely accepted division of psychological processes is the tripartite distinction among action, cognition, and emotion. The theory had a lot to say about the connection between thought and action, but barely mentioned emotion. After developing the theory and devising empirical tests of emergence and optimality, the role that emotion might play in action identification dynamics came into sharp relief. Far from being irrelevant to action identification, emotion plays a key role in calibrating the mind-action connection. It does so by providing the hedonic basis for achieving and maintaining an optimal level of identification.

Emotion serves this function by signaling when mind and action are not well calibrated. In general terms, this contention is consistent with theories suggesting that negative emotion commonly results from the interruption of goal-directed action (cf. Carver and Scheier, 2002; Simon, 1967; Vallacher and Nowak, 1999). Interruption can take many forms, from environmental obstacles to the interference of other people. Our perspective suggests that nonoptimal identification also qualifies as an interrupt – an identity that is either too low- or too high-level disrupts the successful coordination of elements necessary for action implementation. Interruption promotes arousal, which in turn engages epistemic concerns regarding the source of the arousal, presumably in service of choosing an appropriate course of action. In like manner, a breakdown in

mind-action coordination can promote heightened arousal which alerts the performer to possible reasons for the breakdown. In essence, the ineffective performer examines his or her mental content with respect to the action, with an eye toward finding another way to think about what he or she is doing. Arousal stemming from faulty calibration, then, becomes manifest phenomenologically as a special form of self-scrutiny.

Research has established a link between heightened arousal and self-focused attention (Wegner and Giuliano, 1980) and various models concur that attention to the self tends to impair task performance. There is little consensus, however, regarding the means by which self-focused attention produces this effect. In some models, to be self-conscious [p. 337 ↓] is to be aware of the potential self-evaluative implications of what one is doing (e.g., Wicklund and Frey, 1980). Defined in this way, self-focused attention is said to impair performance by reducing attention to task-relevant features. Test anxiety, for instance, is said to occur when the test taker is preoccupied with self-evaluative thoughts (e.g., "I'm going to fail," "I'll be embarrassed") to the relative exclusion of attention to the subtleties of the task at hand (Sarason, 1972; Wine, 1971). But in other formulations (e.g., Kimble and Pelmutter, 1970), self-consciousness refers to heightened awareness of the processes or mechanics underlying the execution of behavior (e.g., the physical movements involved or the coordination of such movements). Models that embrace this definition of self-focused attention argue that a conscious concern with the process of performance essentially disintegrates the action, robbing it of its fluidity and rhythm (cf. Baumeister and Heatherton, 1996).

These contrasting views of self-consciousness map directly onto the optimality hypothesis. The self-evaluative implications of behavior constitute a special class of relatively high-level action identification, whereas the mechanics of behavior reflect considerably lower-level act identities. From the optimality perspective, neither orientation is inherently linked to self-consciousness; rather, each can give rise to such a state depending on the personal difficulty of the action. Specifically, the experience of self-consciousness arises whenever one's conscious level of action control is either too high or too low, given the action's personal difficulty. When performing a simple act or a complex one that has become fairly automated, people will feel self-conscious to the extent that they are conscious of the molecular features of the action. For a difficult act that is best performed with such features in mind, however, self-consciousness is associated instead with sensitivity to the act's larger meanings and effects.

Evidence for this perspective on self-consciousness and its role in performance is provided in the study of speech fluency (Vallacher et al., 1989) discussed above. In addition to tracking participants' speech errors and self-rated performance quality, we asked them to indicate their degree of self-consciousness (as well as anxiety, tension, etc.) while delivering the speech. Results showed that these ratings paralleled the pattern obtained for speech fluency: self-consciousness and related feelings of aversive arousal were greatest when the easy task was identified in low-level terms and the difficult task was identified in high-level terms.

The arousal and negative emotion associated with nonoptimal identification diminishes as people bring their conscious representation of what they are doing into line with the action's difficulty. This does not mean, however, that people's affective state goes flat when they control an action with respect to the optimal level of identification. Rather, people experience positive affect when there is a perfect match between the demands of a task and their mental and behavioral readiness to perform the task. Csikszentmihalyi (1990), for example, has observed that when people experience congruence between their capabilities and the demands of the task – a state he refers to as "flow" – they report diminished self-awareness. The link between optimality and positive emotion is also consistent with research on perceptual fluency (e.g., Winkielman and Cacioppo, 2001), which shows that stimuli are viewed positively to the extent that their features are processed easily. In analogous manner, one can speak of *action fluency*, in which the person has a well-orchestrated and adaptive understanding of an action that facilitates smooth execution of the action. An optimal level of identification enhances the degree to which a person can experience action fluency.

Personality

The optimality hypothesis holds that people will identify their actions at the highest possible [p. 338 ↓] level that still affords effective action control. People clearly differ in their respective expertise in specific domains and thus are also likely to show corresponding differences in the optimal level of identification. A professional tennis player will identify "playing tennis" at a considerably higher level ("winning a match," "improving my ranking") than will (or should) a week-end warrior ("get the ball over the

net," "bend my knees but not too much"). But it also occurred to us that people may differ in their competence more generally, across a variety of action domains. Some people undertake more hobbies during their youth, for example, and become adept at a wider range of activities than do others. There may also be variation in the degree to which individuals encounter information pertaining to the distal consequences of action. Positions of responsibility, for example, may sensitize people to the higher-level implications of action. It could be, too, that learning to appreciate the higher-level identities of action in a few domains creates a readiness to see actions generally in these terms.

We referred to cross-domain differences among people in level of action identification as individual variation in *level of personal agency* (Vallacher and Wegner, 1989). At one end of this dimension is the low-level agent, a person who functions in different domains in a relatively molecular or detailed manner. This person's base-rate tendency is to focus on the mechanistic details of actions. At the other extreme is the high-level agent, someone whose base-rate tendency is to view what he or she does in terms of the action's causal effects, social meanings, and self-evaluative implications. Everyone is likely to have a stable, domain-specific identification level for certain actions, but level of personal agency influences how an individual identifies action across a wide range of domains.

This dimension of individual variation does not represent a trait in the usual sense of the term. A personality trait typically refers to a tendency to emit behaviors from within a content-defined class, such as "sociability" or "conscientiousness." Because any action can be identified in many ways, however, this approach to personality is problematic. "Criticizing an acquaintance," for example, could be viewed as an instance of a trait such as unfriendly, but this classification may miss what the action really meant to the actor. Perhaps he or she was "offering constructive feedback" or "expressing an opinion." Behavioral dispositions reify one particular identity at a relatively high level and thus may fail to capture what people really do. Level of personal agency goes beyond specific behavioral dispositions to address whether the person has trait-like dispositions at all. High-level agents can be expected to enact many of their behaviors under the guidance of higher-level meanings such as self-conceived traits, goals, and values. In contrast, low-level agents tend to engage in actions that are not personally connected to such larger meanings. Level of personal agency, in other words, represents the

degree to which an individual has organized his or her actions into abstract, meaningful categories that can channel behavior into dispositional tendencies.

To explore whether there are cross-domain individual differences in level of action identification, we developed the Behavior Identification Form (BIF) (Vallacher and Wegner, 1989). It consists of 25 act identities, each followed by two alternative identities, one lower and one higher in level. "Making a list," for example, is followed by "Getting organized" (higher-level) and "Writing things down" (lower-level). "Resisting temptation" is followed by "Saying 'no'" (lower-level) and "Showing moral courage" (higher-level). Participants are asked to choose the alternative identity that best describes the action for them. Their level of personal agency is simply the number of high-level identities chosen across the 25 items.

We found that level of personal agency was reliably correlated with several aspects of people's psychology, including action effectiveness, action planning, impulsivity, [p. 339 ↓] action stability, self-monitoring, internal versus external locus of control, and key aspects of self-concept. Compared to low-level agents, the high-level agents were more skilled at a variety of actions (e.g., planning a party, teaching tricks to a pet), had more hobbies, were less impulsive and more intentional in their everyday behavior, were lower in self-monitoring, and had an internal locus of control. High-level agents were also more likely to describe themselves in terms of traits, ascribed greater importance to traits in their self-definition, had higher self-concept certainty, and were less susceptible to social feedback regarding their dispositional qualities. High- and low-level agents did not differ in their level of self-esteem, however, so the relevance of personal agency to certainty and malleability of self-concept is not mediated by the valence of people's self-concept.

Level of personal agency sheds light on the issue of personal versus situation causation, which periodically surfaces as an obsession in social psychology. Low-level agents' behavior tends to be under situational control, in that they enter action contexts with little sense of the action's implications in mind and thus are primed to accept cues to higher-level meaning found in social feedback or situational pressures. High-level agents' behavior, in contrast, tends to be under the control of personal goals and self-conceived tendencies. As a result, they are able to maintain their actions with respect to meaningful representations they carry with them across action contexts. Most people

exist between the extremes of this dimension, so it is not surprising that behavior for people in general reflects a combination of personal and social influences.

Self-Concept

A prevailing wisdom in social psychology is that people will go to great lengths to maintain their self-concept (Tesser et al., 1996), even if the self-concept is unflattering (Baumeister, 1993; Swann, 1990). Thus, people resist new information that might promote a change in their self-perceived qualities, clinging even more tightly to their prevailing self-view. On the other hand, there is reason to believe that self-assessments are strongly impacted by feedback from significant others or even casual acquaintances (cf. Felson, 1989; Mead, 1934). The emergence scenario suggests that both generalizations are valid, but under different circumstances.

A person is likely to deflect or discount social feedback if he or she has a coherent high-level perspective on his or her behavior. If the person knows that he or she is cooperative, for example, he or she is unlikely to embrace feedback suggesting that he or she is really competitive. The self is clearly a familiar object of thought, so the base-rate tendency is to think about one's self-relevant behavior in high-level terms, thereby making self-concept change difficult. But under conditions that promote lower-level identities for an action, a person should show susceptibility to feedback from others that provides an avenue of emergence to higher-level understanding. The emergent identity may be quite different from prior identifications, and thus may provide the person with new "insight" into his or personal make-up.

We investigated the relevance of emergence for self-concept change and stability (Wegner et al., 1986). We arranged for participants to have a "computer analysis" of their personalities. The input for this analysis was participants' description of five things they had done in a recent interaction with someone of the same sex. In the high-level condition, they were asked to describe five things they had done that reflected their opinions, values, and personality traits. In the low-level condition, they were asked to indicate five specific actions reflecting concrete movements and utterances. The computer responded to these inputs with one of two personality profiles – one indicating that the participant was cooperative, the other that he or she was competitive.

Participants were then asked to judge the validity of the computer feedback, describe themselves on various [p. 340 ↓] traits dimensions including cooperative and competitive, and rank order several activities (including both a cooperative activity and a competitive activity) in order of preference for future participation.

By all three measures, results confirmed the emergence scenario. Participants in the high-level condition were skeptical of the computer program and their self-descriptions showed a slight reactance effect – those described as cooperative rated themselves as somewhat competitive and those described as competitive tended to rate themselves as cooperative. This is what one would expect from models that stress self-concept defense (e.g., Swann, 1990). The high-level participants' activity rankings did not show a preference for the activity that reflected the feedback (cooperative or competitive) they had received. High-level identification, then, provided a shield against social feedback. The results were far different for participants in the low-level condition. They judged the computer program to be valid and they rated themselves in accordance with the feedback the program provided – as highly cooperative in response to cooperative feedback but as highly competitive in response to competitive feedback. And their ranking of the future activities reflected their emergent self-under standing. Those provided with cooperative feedback gave a higher ranking to the activity that called for cooperative behavior, but those provided with competitive feedback ranked the competitive activity more favorably.

This scenario of self-concept change has straightforward implications for the effectiveness of psychotherapy intended to change a person's dysfunctional attitudes concerning him or herself. People resist changing their ideas of what they are like, even when these ideas paint a rather gloomy or depressing picture (e.g., Swann et al., 1992). Faced with a person's dysfunctional self-view, a therapist (or a well-intended friend) might be tempted to challenge the person's view directly, encouraging him or her to adopt a more positive attitude. Such a frontal assault may have a temporary impact (e.g., Swann et al., 1990) but is destined to fail or even backfire in relatively short order because it runs counter to the person's coherent and comprehensive self-assessment.

The challenge is to disassemble this high-level identity by getting the person to focus on specific aspects of his or her behavior, and then provide cues to alternative higher-level identities that paint a more flattering self-portrait. Instructing the person to think about

the behavioral evidence for his or her self-assessment, as is done in certain brands of cognitive therapy (e.g., Beck and Weishaar, 2000), trades on this idea. Cognitive behavioral therapy goes a step further by encouraging the client to *do* things rather than simply think (e.g., Meichenbaum, 1977). Because concrete actions require at least some attention to lower-level details, this approach is especially likely to create the low-level mindset that is the precondition for emergent understanding.

Social Influence

Social influence is widely considered to be the pivotal process in social psychology. Any theory worth its salt, then, must have something worthwhile to say about the factors that determine whether or not an individual will change the way he or she thinks or acts with respect to a particular topic or domain. Of course, people can be induced to change their actions and expressed opinions through the application of strong forces, whether rewarding or punitive. But change is likely to be more enduring when it goes beyond enforcing overt behavior to changing the internal dynamics of the target. To promote such changes, it is necessary to disassemble or otherwise destabilize the target's way of thinking, priming him or her for reconfiguration in line with cues to the new message provided by the influence agent.

The disassembly-reconfiguration scenario clearly follows the counters of the emergence process. When people have a coherent [p. 341 ↓] high-level identity for someone's behavior, they are relatively immune to alternative interpretations and evaluations. Metaphorically, people "freeze" when they have a high-level understanding that provides a sense of cognitive closure (cf. Kruglanski and Webster, 1996; Lewin, 1936). But when people identify someone's behavior in relatively low-level terms, they become receptive to coherent perspectives on the behavior provided by social sources and other external factors. In effect, people are motivated to "seize" a higher-level interpretation that provides personal closure concerning the action's meaning.

The extrapolation to social influence is straightforward. The influence agent first induces the target to consider the relevant action in concrete, low-level terms. Simply describing the action in terms of its details can induce low-level identification, as can presenting the target with a surplus of concrete information regarding the action. From this low-

level state, the target experiences a heightened press for coherence. On his or her own, the target might emerge with a higher-level identity that reflects past perspectives or perhaps one that reflects a new integration. But if the influence agent offers a message that provides the missing integration before the target has demonstrated emergence on his or her own, the target may embrace this message as an avenue of emergent understanding, even if it conflicts with his or her prior conception.

The emergence scenario in social influence was tested by asking participants to allocate blame for an alleged rape incident (Vallacher and Selz, 1991). The nature of the incident was such that the motives and intentions of the alleged rapist and victim were open to different interpretations. The incident was presented in the form of a police interview with either the alleged rapist or the victim. Participants read the interview under either a low-level set (reading for detail) or a high-level set (reading for meaning). They then read a police summary concluding either that the perpetrator should be charged with rape or that there were insufficient grounds to press charges. The participants were then asked to allocate responsibility for the event between the perpetrator and the victim. Participants (both males and females) in the low-level condition assigned blame in line with the police summary they read, whereas those who read the interview under a high-level set were not influenced by the police summary. Focusing on “just the facts” in a case of alleged wrongdoing may reduce the influence of one’s personal biases, but this attention to detail can make one all the more vulnerable to influence by other people with biases of their own.

The disassembly-reconfiguration perspective on social influence has been embraced by others in recent years, albeit with considerable refinement and extension to different domains of influence (cf. Knowles and Linn, 2004). Social influence comes in diverse forms (compliance, persuasion, guilt, seduction, etc.), but perhaps these forms are built to a certain extent on a shared platform that reflects people’s press for coherent higher-level understanding (Vallacher et al., 2003).

The Emergent Theory

The extensions we have described were largely unanticipated in the 1980s when we were preoccupied with getting the principles straight. The theory may be poised for yet

further growth, but change this time is likely to take the form of transformation in light of new ideas and methods in psychological science. In particular, two lines of theory and research – *mind perception* and *dynamical social psychology* – represent new ways of framing the dynamics of action identification. The mind perception perspective generalizes action identification principles to the understanding of the minds and actions of other people. It is consistent with several emphases in contemporary psychology, including theory of mind, neural bases of empathy, and the perception (and illusion) of agency. Dynamical [p. 342 ↓] social psychology adapts principles and methods of dynamical systems and complexity, which are at the forefront of contemporary theory and research in the natural sciences, to the investigation of personal and social processes. Within this perspective, the principles of action identification capture a basic dynamic scenario in the mind-action relationship.

Mind Perception

The experience of high-level identification says a lot about the actor's mind. In particular, when the action is complex and extended in time, a high-level mindset suggests both intentionality (acting on purpose, having a plan, working toward a goal) and executive cognitive processes (conscious control of the action, working memory, mindfulness). Actions performed under lower-level identities also require an active mind, of course, but they don't seem to demand the same caliber of mental states. One can move a finger without a great deal of intention and thought, after all, but appreciating the consequences of the finger movement (e.g., sending an e-mail, firing a gun) implicates mental processes rather directly. A recent neuroimaging (fMRI) study has in fact demonstrated greater activation of brain regions associated with higher-order cognitive processes (e.g., temporo-parietal junction) when people are identifying action in high- as opposed to low-level terms (Marsh et al., 2009).

The connection between a person's level of action identification and his or her mental control of the action provided the original focus of the theory. Recently, though, the connection between identification level and mental states has been extended to the perception of other people's actions and minds (Kozak et al., 2006). People are quite willing and able to infer how other people's minds work and such inferences are central to a host of issues in social judgment, including liking, the attribution of personal versus

situational causation, personality judgment, and allocation of responsibility (e.g., Carruthers and Smith, 1996; Epley and Waytz, 2009; Frith and Frith, 2003; Idson and Mischel, 2001; McPherson-Franz and Janoff-Buhnan, 2000; Wegner, 2002). Do people show the same variability in identifying the actions of others as they do in thinking about their own behavior? Is this variability related to the inferences people make about others' mental states? What factors shape each inference – action identification and attribution of mind – and the relationship between such inferences?

To answer these questions, Kozak et al. (2006) modified the BIF (Vallacher and Wegner, 1989) to allow participants to identify a target person's actions rather than their own. They also developed a Mind Attribution Scale (MAS), consisting of ten items assessing participants' inferences about another person's capacity to act with intention, engage in complex cognition, and experience emotion. The BIF and MAS were then employed in several experiments involving vignettes about various hypothetical target persons. Each experiment focused on a particular aspect of the link between action identification, mind attribution, and person perception. Kozak et al. found, first of all, that high-level identification and attribution of mind (intention and complex cognitive processes) were both associated with liking for a target person who was described in fairly neutral terms. Indeed, the level of action identification for liked targets; was often higher than the identification level participants' indicated for their own actions. People are constrained by reality (e.g., personal difficulty, unfamiliarity) in identifying their own actions, but the sky is the limit when thinking about the actions of other people.

By itself, the link between liking and inferences about mental states is open to interpretation; liking may cause high-level identification and mind attribution, but the reverse causal path is also plausible. Kozak et al. (2006) untangled this issue in other experiments by having participants make [p. 343 ↓] judgments about a target person who was clearly likable versus unlikable. The results made clear that evaluation was primary: The liked person was credited with higher-level identities and a more complex mind than the disliked person. This is not cause for concern when likability centers on obvious qualities (e.g., honesty versus dishonesty, sociability versus aloofness). But they found the same relationship when liking was manipulated by the target's fortune versus misfortune. There is evidence that victims of misfortune tend to be derogated by others, even when the misfortune is not of their own doing (Lerner, 1980). When participants read a vignette about a male student in financial straits who could only

afford one meal a day, they reported unfavorable evaluations of him, identified his behavior in lower-level terms, and credited him with less complex cognitions. Victims, it seems, do not have minds like the rest of us.

There is an exception to the connection between liking and high-level identification. Sometimes the people we like do bad things and sometimes the people we don't like do good things. Kozak et al. found that high-level identities were inferred when the valence of the actor matched the valence of the action. So for liked target persons, positive actions were identified at higher levels than were negative actions, but for disliked target persons, negative actions were identified at higher levels. Because high-level identification is linked to personal responsibility, these results are consistent with the tendency to credit liked others (e.g., friends, heroes) for good behavior and to blame disliked others (enemies) for bad behavior. The results are also consistent with research on people's self-presentation of their own actions (Vallacher et al., 1987). People tend to describe their mistakes in terms of lower-level details but to emphasize personal attributes and goals when describing their successes and noteworthy deeds.¹

If higher-level action identification promotes mind perception, it might also undermine processes that lead to the devaluation of minds and, ultimately, to dehumanization. The tendency to see others as animals (Epley and Waytz, 2009), robots (Haslam, 2006), objects (Fredrickson and Roberts, 1997), or as otherwise less than fully human (Harris and Fiske, 2006) may thus depend in part on the deconstruction of mind perception through low-level action identification. It is not yet clear which particular aspects of mind perception are reduced in lower-level action identification – whether, for example, minds are seen as less capable of having experiences or being agentic (Gray et al., 2007), or as less capable of serving as moral agents or moral patients (Gray and Wegner, 2009). What is clear is that identification level influences our perception of minds as worth preserving, and so may ultimately be instrumental in leading people to treat each other as less than human. Action identification may be an initial step toward both the moral regard we give to our most respected conspecifics and as well as to the abject disregard we visit on those we fail to recognize as fellow members of the human race.

Dynamical Social Psychology

The dynamic interplay between higher and lower levels of action identification has a natural resonance with the way dynamical systems in other areas of science evolve, function, and change (cf. Guastello et al., 2009). In its most basic sense, a dynamical system is a set of interconnected elements that influence each other to achieve a common or coordinated state. The resultant higher-order state typically has emergent properties, which simply means that the qualities of the state cannot be reduced to the properties of the constituent elements. Once such a state emerges, it constrains the behavior of the elements that gave rise to it. Because it 'attracts' the system's dynamics, the coherent state is referred to as an *attractor*. A system's attractor stabilizes the system and actively resists change due to outside influences. If change occurs, it is because feedback loops among [p. 344 ↓] the elements are weakened, increasing the degrees of freedom in the system and thereby undermining the coherence of the higher-order state. From this disassembled state of affairs, the system is primed for emergence to a new higher-order state that provides a different configuration of the lower-level elements.

Over the past decade, these hallmarks of a dynamical system have been identified for a variety of social processes, including social judgment, self-concept, social influence, societal transition, and intractable conflict (cf. Vallacher and Nowak, 2007). In self-concept, for example, elements of self-relevant information become integrated to form a coherent perspective on the self, which then constrains the processing of subsequent input, enabling the system to resist change when exposed to inconsistent information or social feedback (Nowak et al., 2000). Self-concept change occurs when the lower-level elements are singled out or decoupled from one another, setting the stage for their reconfiguration with respect to a new and possibly quite different self-view. This basic scenario, which has been detailed as well for the other phenomena indicated above, reflects the essence of the emergence scenario of action identification theory. Not until these manifestations of dynamical social psychology were developed, however, was their genesis in action identification principles appreciated.

Reframing action identification in dynamical terms suggests two refinements of the emergence process. In the basic theory, emergence occurs when people in a low-

level state are provided cues (e.g., social feedback, vivid consequences) that signal an action's higher-level meaning. In a dynamical system, though, emergence can occur without external influence due to self-organization among system elements (Vallacher and Nowak, 1997). This means that the intrinsic dynamics of the mental system can promote emergent meaning – an attractor – for an action's lower-level identities. As a person thinks about and performs a sequence of basic acts, a higher-level identity may spontaneously emerge. Because intrinsic dynamics of mind can take place outside of conscious attention (Port and van Gelder, 1995), new insights into one's action can pop into awareness without warning, in a manner reminiscent of an “Aha!” experience.

The second refinement concerns the potential for multistability in a psychological system (Vallacher and Nowak, 2007). As a system's elements become organized with respect to one coherent state (attractor), the elements that are excluded may form an alternative attractor that can compete for prepotence with the original attractor. When a person develops a highly favorable judgment of someone, for example, inconsistent (i.e., unflattering) information about him or her tends to be discounted. If enough elements of information undergo this fate, they may become organized into an alternative perspective on the target person. If conditions change (e.g., the other person finally goes too far), the judgment system could show a catastrophic shift to the previously latent attractor. The idealized assessment, in other words, could transform into a negative view without going through intermediate steps of disinterest or mild disapproval.

With respect to action, the potential for multistability suggests that although a high-level identity resists change, at some threshold of inconsistent information, the person may suddenly embrace a wholly different high-level identity that has formed by virtue of self-organization among elements that had been discounted in service of maintaining the original identity. A person who stubbornly sees his or her critical comments as constructive despite being told otherwise, for example, might suddenly recognize this behavior as mean-spirited.

The potential for self-organization and multistability have been invoked to understand the nature of social conflicts that have become protracted to the point of seeming intractability, and to suggest new means of resolving such conflicts (e.g., Coleman et al., 2007; Vallacher et al., 2010). When a [p. 345 ↓] conflict develops, the opposing

parties each experience a press for integrative understanding that can provide a coherent and stable platform for action. So although the parties may have a wealth of specific knowledge regarding one another, their respective judgments lose complexity – the separate elements of information become linked by positive feedback loops and take on the same (e.g., negative) higher-level meaning. The resultant coherent state functions as an attractor that incorporates new information and resists external forces that threaten to undermine it.

Some conflict-relevant information, though, cannot be interpreted in line with the attractor. The other party may act in an unambiguously positive manner, for example, or previous positive acts by the party may be made salient in memory. Instances of inconsistent information may be discounted or suppressed when first encountered, but over time they may provide the seeds for an alternative attractor associated with positive thoughts and action possibilities. If conditions should change and promote 'ripeness' for peace, there may be a sudden and dramatic change to this latent attractor. The potential for sudden transitions in the relations between groups mired in conflict has counter-intuitive implications for conflict resolution. Rather than addressing the issues that launched the conflict, a more effective strategy is to create the basis for an alternative way of thinking and behaving that is likely to be dismissed in the short run but which creates an alternative coherent state that can become manifest in the long run.

The Metatheory

At the most basic level, action identification theory is a set of principles concerning the representation and control of action. The principles capture the conflicting forces that interact to promote a particular form of understanding in the face of a multitude of equally plausible identities for one's actions. This basic identity for the theory gives rise to a number of implications that expand the possible ways in which the theory can be seen. The interplay of the theory's principles is manifest in several notable phenomena, including emotion, stability versus change in self-concept, social judgment, social influence, and individual variation in the mind-action relationship. Which of these aspects of the theory are prepotent depends on the social context surrounding the lay person and the research agenda of the psychologist.

At a yet higher level of understanding, action identification theory can be looked upon as a basic dynamic for mind and action that defines human experience. The theory captures the interplay between the often-competing concerns with comprehensive understanding and effective action that underlies personal functioning across social contexts. Because life itself is dynamic, this interplay is iterated continuously on different (embedded) timescales, ensuring complexity and growth as people go about their daily lives. Action identification, in this light, is a specific, lower-level manifestation of a pervasive dynamic that coordinates the interplay between mind and reality in people's lives. Our goal in this chapter was to illustrate this dynamic. Of course, we were also focused on keystrokes and coming up with a concluding thought. We just did both.

Notes

1 The results are also consistent with research on intergroup biases in the language used to describe action (Maass et al., 1995). People describe the positive behaviors of ingroup members at more abstract levels than they do the positive behaviors of out-group members, but describe the negative behaviors of outgroup members at more abstract levels than they do the negative behaviors of ingroup behaviors. This resemblance should be viewed cautiously, though, because identification levels center on means-ends relations, not on levels of abstraction per se. For example, “pushing a button” is higher-level than “moving a finger” by virtue of the *by* [p. 346 ↓] relation, but the former does not seem to be all that more abstract than the latter.

References

Baumeister, R.F. (ed.) (1993) Self-esteem: The Puzzle of Low Self-regard . New York: Plenum. <http://dx.doi.org/10.1007/978-1-4684-8956-9>

Baumeister, R.F. and Heatherton, T.F. (1996) Self-regulation failure: An overview . Psychological Inquiry , 7, 1–15. http://dx.doi.org/10.1207/s15327965pli0701_1

Beck, A.T. and Weishaar, M. (2000) Cognitive therapy . In R.J. Corsini, ed. and D. Wedding (eds), *Current Psychotherapies* , 6th Edition, pp. 241–272. Itasca, IL: F.E. Peacock.

Carver, C.S. and Scheier, M.F. (1981) Attention and Self-regulation: A Control-theory Approach to Human Behavior . New York: Springer-Verlag. <http://dx.doi.org/10.1007/978-1-4612-5887-2>

Carver, C.S. and Scheier, M.F. (2002) Control processes and self-organization as complementary principles underlying behavior . *Personality and Social Psychology Review* , 6, 304–315. http://dx.doi.org/10.1207/S15327957PSPR0604_05

Carruthers, P., ed. and Smith, P.K. (eds) (1996) *Theories of Theories of Mind* . Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511597985>

Coleman, P.T., Vallacher, R.R., Nowak, A. and Bui-Wrzosinska, L. (2007) Intractable conflict as an attractor: A dynamical systems approach to conflict escalation and intractability . *American Behavioral Scientist* , 50, 1454–1475. <http://dx.doi.org/10.1177/0002764207302463>

Csikszentmihalyi, M. (1990) *Flow: The Psychology of Optimal Experience* . New York: Harper & Row.

Danto, A. (1963) What we can do . *Journal of Philosophy* , 40, 435–445. <http://dx.doi.org/10.2307/2023429>

Epley, N. and Waytz, A. (2009) Mind perception . In D.T. Gilbert, ed. , and S.T. Fiske (eds), *The Handbook of Social Psychology* , 5th Edition. New York: Oxford University Press.

Felson, R.B. (1989) Parents and the reflected appraisal process: A longitudinal analysis . *Journal of Personality and Social Psychology* , 56, 965–971. <http://dx.doi.org/10.1037/0022-3514.56.6.965>

Fredrickson, B.L. and Roberts, T.A. (1997) Objectification theory . *Psychology of Women Quarterly* , 21, 173–206. <http://dx.doi.org/10.1111/j.1471-6402.1997.tb00108.x>

Frith, U. and Frith, C.D. (2003) Development and neurophysiology of mentalizing . Philosophical Transactions: Biological Sciences , 358, 459–473. <http://dx.doi.org/10.1098/rstb.2002.1218>

Gergen, K.J. (1985) The social constructionist movement in modern psychology . American Psychologist , 40, 266–275. <http://dx.doi.org/10.1037/0003-066X.40.3.266>

Goldman, A.I. (1970) A Theory of Human Action . Princeton, NJ: Princeton University Press.

Gray, H.M., Gray, K. and Wegner, D.M. (2007) Dimensions of mind perception . Science , 315, 619. <http://dx.doi.org/10.1126/science.1134475>

Gray, K. and Wegner, D.M. (2009) Moral typecasting: Divergent perceptions of moral agents and moral patients . Journal of Personality and Social Psychology , 96, 505–520. <http://dx.doi.org/10.1037/a0013748>

Guastello, S., ed. , Koopmans, M., ed. and Pincus, D. (eds) (2009) Chaos and Complexity in Psychology: The Theory of Nonlinear Dynamical Systems . New York: Cambridge University Press.

Hampshire, S. (1960) Thought and Action . London: Chatto and Windus.

Harris, L.T. and Fiske, S.T. (2006) Dehumanizing the lowest of the low: Neuro-imaging responses to extreme outgroups . Psychological Science , 17, 847–853. <http://dx.doi.org/10.1111/j.1467-9280.2006.01793.x>

Haslam, N. (2006) Dehumanization: An integrated review . Personality and Social Psychology Review , 10, 252–264. http://dx.doi.org/10.1207/s15327957pspr1003_4

Higgins, E.T. (1998) Promotion and prevention: Regulatory focus as a motivational principle . In M.P. Zanna (ed.), Advances in Experimental Social Psychology Vol . 30, 1–46. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960381-0>

Idson, L.C and Mischel, W. (2001) The personality of familiar and significant people: The lay perceiver as a social-cognitive theorist . *Journal of Personality and Social Psychology* , 80, 585–596. <http://dx.doi.org/10.1037/0022-3514.80.4.585>

James, W. (1890) *Principles of Psychology* . New York: Holt. <http://dx.doi.org/10.1037/11059-000>

Kimble, G.A. and Perlmutter, L.C. (1970) The problem of volition . *Psychological Review* , 77, 361–384. <http://dx.doi.org/10.1037/h0029782>

Knowles, E.S., ed. and Linn, J.A. (eds) (2004) *Resistance and Persuasion* . Mahwah, NJ: Erlbaum.

Kozak, M.N., Marsh, A.A. and Wegner, D.M. (2006) What do I think you're doing? Action identification and mind attribution . *Journal of Personality and Social Psychology* , 90, 543–555. <http://dx.doi.org/10.1037/0022-3514.90.4.543>

Kruglanski, A.W. (1975) The endogenous-exogenous partition in attribution theory . *Psychological Review* , 82, 387–406. <http://dx.doi.org/10.1037/0033-295X.82.6.387>

Kruglanski, A.W. and Webster, D.M. (1996) Motivated closing of the mind: ‘Seizing’ and ‘freezing’ . *Psychological Review* , 103, 263–283. <http://dx.doi.org/10.1037/0033-295X.103.2.263>

Langer, E.J. and Imber, L.G. (1979) When practice makes imperfect: Debilitating effects of overlearning . *Journal of Personality and Social Psychology* , 37, 2014–2024. <http://dx.doi.org/10.1037/0022-3514.37.11.2014>

Lerner, M.J. (1980) *The Belief in a Just World: A Fundamental Delusion* . New York: Plenum.

Lewin, K. (1936) *Principles of Topological Psychology* . New York: McGraw-Hill. <http://dx.doi.org/10.1037/10019-000>

Maass, A., Milesi, A., Zabbini, S. and Stahlberg, D. (1995) Linguistic intergroup bias: Differential expectancies or in-group protection? *Journal of Personality and Social Psychology* , 68, 116–126. <http://dx.doi.org/10.1037/0022-3514.68.1.116>

Marsh, A.A., Kozak, M.N., Reid, M., Yu, H., Wegner, D.M. and Blair, R.J.R. (2009) The neural substrates of action identification . Unpublished manuscript, Georgetown University.

McPherson-Franz, C. and Janoff-Bulman, R. (2000) Considering both sides: The limits of perspective-taking . *Basic and Applied Social Psychology* , 22, 31–42.

Mead, G.H. (1934) *Mind, Self, and Society* . Chicago: University of Chicago Press.

Meichenbaum, D. (1977) *Cognitive Behavior Modification* . New York: Plenum.

Miller, G.A., Galanter, E. and Pribram, K.H. (1960) *Plans and the Structure of Behavior* . New York: Holt. <http://dx.doi.org/10.1037/10039-000>

Newtonson, D. (1976) Foundations of attribution: The perception of ongoing behavior . In J.H. Harvey, ed. , W.J. Ickes, ed. , and R.F. Kidd (eds), *New Directions in Attribution Research* , 1, 223–247. Hillsdale, NJ: Erlbaum.

Nowak, A., Vallacher, R.R., Tesser, A. and Borkowski, W. (2000) Society of self: The emergence of collective properties in self-structure . *Psychological Review* , 107, 39–61. <http://dx.doi.org/10.1037/0033-295X.107.1.39>

Port, R.F., ed. and van Gelder, T. (eds) (1995) *Mind as Motion: Explorations in the Dynamics of Cognition* . Cambridge, MA: MIT Press.

Ryle, G. (1949) *The Concept of Mind* . London: Hutchinson.

Sarason, I.G. (1972). Experimental approaches to test anxiety: Attention and the uses of information. In C.D. Spielberger (ed.), *Anxiety: Current Trends in Theory and Research* , 2, 383–403. New York: Academic Press.

Simon, H.A. (1967) Motivational and emotional controls of cognition . Psychological Review , 74, 29–39. <http://dx.doi.org/10.1037/h0024127>

Swann, W.B., Jr. (1990) To be adored or to be known? The interplay of self-enhancement and self-verification . In E.T. Higgins, ed. and R.M. Sorrentino (eds), Handbook of Motivation and Cognition , 2, 408–448. New York: Guilford.

Swann, W.B., Jr., Hixon, J.G., Stein-Seroussi, A. and Gilbert, D. (1990) The fleeting gleam of praise: Cognitive processes underlying behavioral reactions to self-relevant feedback . Journal of Personality and Social Psychology , 59, 17–26. <http://dx.doi.org/10.1037/0022-3514.59.1.17>

Swann, W.B., Jr., Wenzlaff, R.M., Krull, D.S. and Pelham, B.W. (1992) Allure of negative feedback: Self-verification strivings among depressed persons . Journal of Abnormal Psychology , 101, 293–306. <http://dx.doi.org/10.1037/0021-843X.101.2.293>

Tesser, A., Martin, L. and Cornell, D. (1996) On the substitutability of self-protective mechanisms . In P.M. Gollwitzer, ed. and J.A. Bargh (eds), The Psychology of Action , pp. 48–68. New York: Guilford.

Trope, Y. and Lieberman, N. (2003) Temporal construal . Psychological Review , 110, 403–421. <http://dx.doi.org/10.1037/0033-295X.110.3.403>

Vallacher, R.R. and Nowak, A. (1997) The emergence of dynamical social psychology . Psychological Inquiry , 4, 73–99. http://dx.doi.org/10.1207/s15327965pli0802_1

Vallacher, R.R. and Nowak, A. (1999) The dynamics of self-regulation . In R.S. Wyer, Jr. (ed.), Advances in Self-regulation , 12, 241–259. Mahwah, NJ: Erlbaum.

Vallacher, R.R. and Nowak, A. (2007) Dynamical social psychology: Finding order in the flow of human experience . In A.W. Kruglanski, ed. and E.T. Higgins (eds), Social Psychology: Handbook of Basic Principles , pp. 734–758. New York: Guilford.

Vallacher, R.R., Nowak, A. and Miller, M.E. (2003) Social influence and group dynamics . In T. Millon, ed. and M.J. Lerner (eds), Handbook of Psychology, Vol. 5: Personality and Social Psychology , 383–417. New York: Wiley.

Vallacher, R.R. and Selz, K. (1991) Who's to blame? Action identification in allocating responsibility for alleged rape . *Social Cognition* , 9, 194–219. <http://dx.doi.org/10.1521/soco.1991.9.2.194>

Vallacher, R.R. and Wegner, D.M. (1985) A Theory of Action Identification . Hillsdale, NJ: Erlbaum.

Vallacher, R.R. and Wegner, D.M. (1987) What do people think they're doing? Action identification and human behavior . *Psychological Review* , 94, 3–15. <http://dx.doi.org/10.1037/0033-295X.94.1.3>

Vallacher, R.R. and Wegner, D.M. (1989) Levels of personal agency: Individual variation in action identification . *Journal of Personality and Social Psychology* , 57, 660–671. <http://dx.doi.org/10.1037/0022-3514.57.4.660>

Vallacher, R.R., Wegner, D.M. and Frederick, J. (1987) The presentation of self through action identification . *Social Cognition* , 5, 301–322. <http://dx.doi.org/10.1521/soco.1987.5.3.301>

Vallacher, R.R., Wegner, D.M. and Somoza, M.P. (1989) That's easy for you to say: Action identification and speech fluency . *Journal of Personality and Social Psychology* , 56, 199–208. <http://dx.doi.org/10.1037/0022-3514.56.2.199>

Wegner, D.M. (2002) The illusion of Conscious Will . Cambridge, MA: MIT Press.

Wegner, D.M. and Giuliano, T. (1980) Arousal-induced attention to self . *Journal of Personality and Social Psychology* , 38, 719–726. <http://dx.doi.org/10.1037/0022-3514.38.5.719>

Wegner, D.M. and Vallacher, R.R. (1986) Action identification . In R.M. Sorrentino, ed. and E.T. Higgins (eds), *Handbook of Motivation and Cognition* , 1, 550–582. New York: Guilford.

Wegner, D.M., Vallacher, R.R., Kiersted, G. and Dizadji, D. (1986) Action identification in the emergence of social behavior . *Social Cognition* , 4, 18–38. <http://dx.doi.org/10.1521/soco.1986.4.1.18>

Wegner, D.M., Vallacher, R.R., Macomber, G., Wood, R. and Arps, K. (1984) The emergence of action . *Journal of Personality and Social Psychology* , 46, 269–279.
<http://dx.doi.org/10.1037/0022-3514.46.2.269>

Wicklund, R.A. and Frey, D. (1980) Self-awareness theory: When the self makes a difference . In D.M. Wegner, ed. and R.R. Vallacher (eds), *The Self in Social Psychology* pp. 31–54. New York: Oxford University Press.

Wine, J.D. (1971) Test anxiety and the direction of attention . *Psychological Bulletin* , 76, 92–104. <http://dx.doi.org/10.1037/h0031332>

Winkielman, P. and Cacioppo, J.T. (2001) Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation leads to positive affect . *Journal of Personality and Social Psychology* , 81, 989–1000. <http://dx.doi.org/10.1037/0022-3514.81.6.989>

<http://dx.doi.org/10.4135/9781446249215.n17>

Handbook of Theories of Social Psychology: Volume 1

Social Cognitive Theory

Contributors: Albert Bandura

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Social Cognitive Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n18>

Print pages: 349-375

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n18>

[p. 349 ↓]

Chapter 17: Social Cognitive Theory

Albert Bandura, ed.

Abstract

The present chapter traces the evaluation of social cognitive theory toward a psychology of human agency. The nature of human agency and the mechanisms through which it operates is analyzed in the context of major changes in the conception of human behavior.

These theoretical orientations include behaviorism, the tight hold of the psychoanalytic movement on psychology and the pop culture, conceptions of mind as symbol manipulator in the likeness of the linear computer, eliminative physicalism contending that human behavior is shaped and regulated by neurophysiological mechanisms that operate outside one's awareness and control, and the growing ascendancy of human agency in the coevolution process. The utility of a psychological theory is judged by three criteria: its explanatory power, its predictive power and, in the final analysis, its operative power to effect personal and social change. Social cognitive theory lends itself readily to social applications. Three illustrative applications document the transformative changes in the field of psychotherapy, development of large-scale health promotion systems founded on a shift in the health field from a disease model to a health model, and global applications that address some of the most urgent worldwide problems.

Introduction

The present chapter traces the evolution of social cognitive theory toward a psychology of human agency. To be an agent is to influence the course of events by one's actions. In this view, people are contributors to their life circumstances, not just products of them (Bandura, 2006a, 2008a). The agentic theoretical perspective serves as the integrative principle in human self-development, adaptations, and change. Human functioning is rooted in social systems. Therefore, personal agency operates within a broad network of sociostructural influences. In these agentic transactions, people create social systems and the practices of social systems, in turn, influence personal development and functioning. Given this dynamic bidirectional influence, social cognitive theory rejects a dualism between personal agency and a social structure disembodied from human activity.

The theoretical framework guiding my work was originally labeled *social learning theory*. [p. 350 ↓] I later relabeled the theory as *social cognitive theory* for several reasons (Bandura, 1986). A variety of theories founded on divergent tenets – Miller and Dollard's drive theory, Rotter's outcome expectancy theory, Gewirtz's operant conditioning theory, and Patterson's functionalist theory – were all christened social learning theory. This created untold confusion in the literature concerning the theory being tested. Moreover, the theory under discussion had always been much broader than the initial descriptive label as a theory of learning. It not only addressed how people acquire knowledge and competencies but also how they motivate and regulate their behavior and create social systems that organize and structure their lives. In the more fitting appellation as *social cognitive theory*, the *social* portion of the title acknowledges the social origins of much human thought and action; the *cognitive* portion recognizes the influential contribution of cognitive processes to human motivation, affect, and action.

When I began my career, behaviorism had a stranglehold on the field of psychology. It focused primarily on learning by direct experience through paired stimulation and rewarding and punishing response consequences. Methodological reductionism prescribed the research strategy, typically with rats and pigeons, on the assumption that understanding rudimentary learning processes would explain complex human behavior.

Cognitive processes were dismissed as redundant inner links in the chain of causation or as explanatory fictions.

The behavioristic theorizing was discordant with the evident social reality that much of what we learn is through the power of social modeling. I found it difficult to conceive of a culture in which its intricate competencies, language, mores, customs, and familial, educational, occupational, religious, and political practices were laboriously shaped by rewarding and punishing consequences of trial-and-error performances. This tedious and potentially hazardous process is short cut by social modeling.

Centrality and Pervasiveness of Social Modeling

Despite the centrality and pervasiveness of social modeling there was little theorizing and research on the nature, scope, and mechanisms governing this basic mode of social influence. Some of the early accounts of social modeling characterized it as imitation and marginalized it as simply mimicry of specific acts. This narrow conception discouraged interest in the phenomenon. For example, Miller and Dollard (1941), viewing social modeling from the behavioristic perspective, treated it as a special case of discrimination learning. A model provides a social cue, the observer performs a matching response, and its reinforcement strengthens the tendency to behave imitatively. Personality and developmental theorists conceptualized it as identification involving wholesale incorporation of modeled attributes. The defining properties of identification were too diffuse, arbitrary, and empirically questionable either to clarify modeling processes or to guide scientific inquiry (Bandura, 1969). I found these early conceptions seriously wanting on the determinants, mechanisms, and scope of social modeling. We launched a program of research on social modeling as it typically occurs observationally in the absence of reinforced performance.

In a chapter entitled “Vicarious Processes: A Case of No-trial Learning” (Bandura, 1965), I presented the findings of our studies showing that observational learning through exposure to models requires neither response enactment nor reinforcement. Social modeling operated through four cognitive sub-functions encompassing

attentional, representational, enactive translational, and motivational processes (Bandura, 1971a). I came under heavy fire from operant conditioners for whom nonreinforced modeling posed a major problem for their explanatory system (Baer et al., 1967). They contended that reinforcement of some matching responses established imitation as a conditioned reinforcer. We demonstrated that generalized imitation is governed by beliefs [p. 351 ↓] about how to influence the models' behavior and outcome expectations for modeled activities rather than by conditioned reinforcers.

Social cognitive theory broadened the scope of modeling influences and the functions it serves (Bandura, 1986). In addition to cultivating cognitive and behavioral competencies, modeling influences were shown to alter motivation, create and modify emotional proclivities, serve as social prompts that activate, channel, and support given styles of behavior, and shape images of reality.

Power and Reach of Symbolic Modeling

A growing influential source of social learning is the varied and pervasive symbolic modeling through the electronic media. A major advantage of symbolic modeling is that it can transmit information of virtually limitless variety to vast populations simultaneously in widely dispersed locales. Indeed, the extraordinary advances in the technology of communication are transforming the nature, reach, speed, and loci of human influence (Bandura, 2002a). Modeled new ideas, values, and styles of conduct are now being rapidly spread worldwide in ways that foster a globally distributed consciousness.

Social cognitive theory addressed the personal and social structural factors that determine the adoption of innovation and the social networks through which the influence diffuses (Bandura, 2006b). The evolving information technologies increasingly serve as a vehicle for building social networks that transcend the barriers of time and space. The Internet provides instant communicative access worldwide. Global broadcasts now show sociopolitical conflicts, the strategies and countermeasures used, and their effects as they are happening. This makes electronic modeling a powerful vehicle for transcultural and sociopolitical change (Bandura, 2002a; Braithwaite, 1994).

Our traditional theories of human behavior were formulated long before these revolutionary advances in communication technologies. The theoretical explanations of human behavior were heavily rooted in influences from the local social environment and the effects of enactive experiences. The social reality of contemporary society is markedly different with growing ascendancy of the symbolic environment and increased opportunities for the exercise of personal and collective agency in self-development, adaptation, and change.

Correcting Misconceptions about the Nature and Scope of Modeling

There were a number of entrenched misconceptions about the nature and scope of modeling that put a damper on research and social applications of this powerful mode of learning and social influence. Progress in this area, therefore, required research designed not only to elucidate the determinants and mechanisms of social modeling but also to put the misconceptions to rest. One such misconception was that modeling, construed as “imitation,” could produce only response mimicry. This is largely the legacy of the early narrow conceptions of modeling. Exemplars usually differ in content and other details but embody the same underlying principle. To cite a simple example, the passive linguistic form may be embodied in any variety of utterances. Research on abstract modeling showed that modeling involves abstracting the information conveyed by specific exemplars about the structure and the underlying principles governing the behavior, rather than simply mimicking the specific exemplars (Bandura, 1986; Rosenthal and Zimmerman, 1978). Once individuals learn the guiding principle, they can use it to generate new versions of the behavior that go beyond what they have seen or heard. They can tailor the behavior to suit changing circumstances.

There was another oft-repeated misconception regarding the scope of modeling. Many activities [p. 352 ↓] involve cognitive skills on how to acquire and use information for predicting and solving problems. Critics argued that modeling cannot build cognitive skills because thought processes are covert and are not adequately reflected in modeled actions, which are the end-products of the cognitive operations. This was a limitation of conceptual vision rather than an inherent limitation of modeling. Cognitive

skills can be readily exemplified and cultivated by cognitive modeling in which models verbalize aloud their reasoning strategies as they engage in problem-solving activities. The thoughts guiding their decisions and actions are thus made observable. Cognitive modeling was shown to be more powerful in enhancing perceived self-efficacy and building complex cognitive skills than the commonly used tutorial methods.

Another misconception requiring retirement held that modeling is antithetical to creativity. We were able to show how innovation can emerge through modeling. When exposed to models who differ in their styles of thinking and behavior, observers rarely pattern their behavior exclusively after a single source. Nor do they adopt all the attributes even of preferred models. Rather, observers combine various features of different models into new amalgams that differ from the individual modeled sources. Thus, two observers can construct new forms of behavior solely through modeling that differ from each other by selectively blending different features from the variant models. In many social and technological innovations, individuals adopt modeled aspects found to be effective, improve upon them, synthesize them into new forms, and tailor them to their particular circumstances. In short, selective modeling is often the mother of innovation.

Baptism in Power Politics

At the time that I was conducting the modeling experiments in the late 1950s, television had diffused rapidly throughout society. The advent of television vastly expanded the range of models available to the general public. The broadcast industry traded heavily on gratuitous violence in the belief that violence sells. Television provided viewers with unlimited opportunities day in and day out to learn the whole gamut of homicidal conduct within the comfort of their homes. There was growing public concern about the possible effects of televised violence on children.

Among the different experimental methods I used to study observational learning was the oft-cited Bobo doll studies on the acquisition of novel forms of aggression through modeling (Bandura et al., 1963) The theory in vogue at the time contended that exposure to modeled aggression is cathartic. It reduces aggression by draining aggressive impulses. We found otherwise. Children who had observed an adult aggress

in unique verbal and physical ways toward an inflated Bobo doll modeled the aggressive styles of conduct. They also were less restrained in expressing, in their play activity, aggression they had learned elsewhere, such as attraction to guns.

I was invited to testify before the Senate Communications Committee, the Federal Trade Commission, and the National Commission on the Causes and Prevention of Violence prompted by the assassination of Robert Kennedy. The Federal Trade Commission was troubled by increasing reports of serious injuries suffered by children who modeled hazardous activities in televised advertisements. The commission used our research findings on modeling to get advertisers to alter ads depicting injurious feats by children on bicycles and dune buggies, ads for headache remedies in which the characters induce splitting headaches by pounding each other on the head with mallets, and other types of ads showing children performing activities that risk serious injury.

This excursion into the public policy arena provided a sobering glimpse into the power of the broadcast industry, some of which was directed at me personally. I got my first inkling into the exercise of this power at a [p. 353 ↓] meeting convened by the National Institute of Mental Health (NIMH) to draft a research agenda on television's effects. Surprisingly, we met at the plush Waldorf Towers in New York rather than in Washington for what turned out to be essentially a production staged by the broadcast industry under the auspices of NIMH. After we identified the different lines of research that could advance the understanding of television's effects, the research community was invited to submit grant proposals. A review panel, meeting in a luxurious Caribbean setting, rejected my proposal.

Look magazine invited me to write a piece on the social influence of television for a special issue they were publishing on youth. When it appeared, the Television Information Office, a subsidiary of the National Association of Broadcasters, sent a large packet to its sponsor stations explaining why my research on social modeling should be disregarded. This was just the beginning of a multipronged offense. Psychologist Ruth Hartley prepared a document commissioned by CBS in which she took me to task and criticized the relevance of other experimental studies demonstrating a positive relation between exposure to violent fare and aggressive behavior. In an editorial in *TV Guide* titled, "A Child is Not Rat," Edith Efron (1969a), senior editor of *TV Guide*, misrepresented the modeling studies as "conditioning studies." She cited Hartley,

whom she dubbed the “best-known attacker,” as the authoritative critic of experimental research on the effects of televised violence. In an expansive indictment, Efron not only took issue with my study, but included “virtually all of his colleagues” as well in the critique of experimental studies of aggressive modeling.

With financial sponsorship and coproduction by CBS, Milgram and Shotland (1973) conducted studies showing that exposure to modeled thievery does not lead viewers to steal money from a charity box labeled *Good Ship Hope* for a medical charity that treats poor children worldwide. The charity box was mounted on a poster showing a physician treating a little girl and a picture of the hospital ship with the words, “Where there is Hope there is life.” This experimental setup is analogous to demonstrating television null effects by showing that viewers will not rip off charitable contributions to Mother Theresa. As the saying goes, there is honor even among thieves. The studies were published as a book and distributed free of charge by the network. They did not survive conceptual and empirical scrutiny. In an author editorial in *TV Guide* under the title “The Man in the Eye of the Hurricane,” Edith Efron (1969b) dismissed the modeling studies, complained that the research by members of the “Bandura school … won them center stage in Washington,” and criticized the Surgeon General’s office for acting “as if Rome were burning and Dr. Bandura were a fire extinguisher” (1969: 37).

One evening I received a call from one of my graduate students telling me to turn on my television set to see the character playing my role undergoing a blistering cross-examination concerning the modeling studies. I wasn’t doing too well! In the plot-line of this televised movie, a beleaguered wife of a screenwriter defends him as he is being unmercifully victimized by a haranguing press and a vindictive mother who claims her son’s crime was prompted by a similar act in one of the screenwriter’s televised plots. The cross-examiner was disputing evidence that televised violence affects aggressive behavior. As I was being pummeled by media-commissioned critiques, sponsored studies, paid consultants, and fictionalized dramas, I began to feel a kinship with the battered Bobo doll!

Failure to distinguish between the diverse effects of televised violence and the appropriate methodologies for elucidating them provided a fertile ground for disputes. Different lines of research identified four major effects of exposure to televised violence: it can teach novel aggressive styles of conduct; weaken restraints over interpersonal

aggression by legitimizing, glamorizing, and [p. 354 ↓] trivializing violent conduct; desensitize and habituate viewers to human cruelty; and shape public images of reality. Each of these separable effects requires a different methodology (Bandura, 2009).

I had to address misunderstandings and misrepresentations of the research using novel modeled aggression to study observational learning. The mistaken critique, which continues to be repeated in our textbooks, is that the study used a nonhuman target and Bobo dolls are for punching. The Bobo doll laboratory experiments were designed to clarify the processes governing observational learning. The methodology for measuring learning effects requires conditions in which viewers feel free to reveal all they have learned. This requires simulated targets rather than retaliative ones. To use human targets to assess the instructive function of televised influence would be as nonsensical as to require bombardiers to bomb San Francisco, New York, or some other inhabited locations to test the extent to which they had acquired bombing skills.

We were not interested in whether children punched the Bobo doll. Rather, we measured whether children assaulted it in the novel modeled ways, such as pummeling it with a mallet and voiced the novel aggressive neologisms as they assaulted the doll. Children in the control condition never exhibited the highly novel form of aggression. Although modeled aggression was only one among a variety of experimental paradigms we used to clarify the mechanisms governing diverse modeling effects, it is the one that is featured in portrayals of social cognitive theory.

There are more chapters to the exercise of political leverage regarding research on media effects. The National Commission on the Causes and Prevention of Violence (1969) was about to release its report concluding, in the mass media section, that the empirical evidence taken as a whole was supportive of a positive relation between televised violence and aggressive behavior. In a surprise move, Senator Pastore, a supporter of the broadcast industry (Paisley, 1972) who chaired the Communications Subcommittee, instructed the Surgeon General, with President Nixon's endorsement, to assemble a committee of experts to evaluate the effects of televised violence and to allocate a million dollars for new research on this topic. The first meeting of the evaluation committee took place at the Center for Advanced Study at Stanford. Ed Parker, who coauthored a book on *Television in the Lives of Children*, and I were invited to sit in on the meeting. We were surprised to find that 40 percent of the committee

membership were tied to the broadcast industry – two network researchers, two network consultants, and a former research executive at CBS.

We enlisted Senator Metcalf, a Stanford graduate, to obtain information on the selection procedure. Health, Education, and Welfare Secretary Finch explained that each network was allowed to veto, without explanation, any of the nominees on the list submitted by professional associations and the broadcast networks. I was one of eight researchers, including Len Berkowitz, Percy Tannenbaum, child psychiatrist Leon Eisenberg, and sociologists Leo Bogart and Otto Larsen, who were vetoed. Finch provided two justifications for the veto procedure – precedent and objectivity. He explained that the tobacco industry was given veto power in the formation of the committee to evaluate the health effects of smoking. The media report would have greater impact, he claimed, if the committee members were entirely objective. Senator Metcalf was astonished to learn that the tobacco industry was also given sole veto power. He questioned the selective privilege of veto power given to the broadcast industry and how stacking the committee with folks tied to the television industry accomplished impartiality.

Writing the report created headaches for the broadcast-linked members because the empirical data were not friendly to the conclusion of cathartic or null effects. The report was written in opaque technobabble that was better suited to confuse than to inform the public. Rose Goldsen (1972), a Cornell sociologist, dubbed the report “science in wonderland.” Before the report was released, a copy was leaked to Jack Gould (1972) of the *New York Times*, which published a column [p. 355 ↓] on the report under the misleading headline, “TV Violence Held Unharmful to Youth.”

Researchers who conducted the studies for the Committee were incensed at the misrepresentation of their findings. They protested to Senator Pastore, who then scheduled an open Senate hearing on the committee's report. After years of obfuscation, negation, and disparagement of research programs by the broadcast industry, their own chief researcher, Joseph Klapper, acknowledged at the hearings, “There were indications of a causal relationship … The catharsis theory had no empirical support.” No US network reported on the Senate hearing. Because of concern over the spillover of US televised violence into Canada, the Film Board of Canada (1972) filmed the entire Senate hearing. Several social scientists reported on the perversion of the scientific review process. Mathilda Paisley (1972) wrote a piece on

violence done to TV violence research. In a book devoted to this controversial episode, *TV Violence and the Child*, Cater and Strickland (1975) traced the evolution and fate of the report. *Science* published a lead article documenting and condemning the misuse of the scientific advisory system for policy initiatives (Boffey and Walsh, 1970).

The late President Johnson once remarked that politics is like sausage making. You don't want to examine what goes into it. Social scientists seek to advance knowledge that can inform public policy. As revealed in the stealthy workings of the sociopolitical forces swirling around the issue of television violence, we also need to study how politics and power, which shape public policy, determine how our knowledge is used. Policy research is difficult to conduct, and we do little of it.

Transformation of the Field of Psychotherapy

While behaviorism ruled over general psychology when I entered the field, psychodynamic theory, especially the psychoanalytic form, reigned over the fields of personality, developmental, psychotherapy, and the pop culture. The mid-1950s witnessed growing disillusionment with this line of theorizing and its mode of treatment. The theory lacked predictive power and did not fare well in therapeutic effectiveness. Dick Walters and I provided an alternative view of human behavior in the book, *Social Learning and Personality Development* (Bandura and Walters, 1963).

During this period, I was teaching courses on personal and social change at Stanford. I was intrigued by cases in which direct modification of problem behavior not only produced lasting improvements in people's lives but fostered generalized benefits in nontreated areas of functioning. Drawing on an emerging literature on psychosocial change, I published the article "Psychotherapy as a Learning Process" in the *Psychological Bulletin* (Bandura, 1961). It was organized around six basic principles of personal change.

The time was apparently ripe for a new direction in the conceptualization and treatment of behavior. I was flooded with reprint requests from home and abroad across

specialties and disciplinary domains. Eysenck invited me to contribute a chapter to a volume he was editing. The chapter kept enlarging until it outgrew the assignment. Instead, it turned into the volume *Principles of Behavior Modification* (Bandura, 1969). It addressed the influential role of cognitive, vicarious, and self-regulatory mechanisms in human functioning.

We were devising new modes of treatment for phobic conditions using guided mastery experiences as the principal vehicle of change. With appropriate mastery aids seemingly unachievable changes become doable. (Bandura et al., 1969). This proved to be a consistently powerful treatment that instilled a robust sense of coping efficacy; transformed attitudes toward the phobic objects from abhorrence to liking and wiped out anxiety, biological stress reactions, and phobic behavior. These people had been [p. 356 ↓] plagued by recurrent nightmares for 20 or 30 years. One of the most striking changes was the power of mastery experiences to transform dream activity and wipe out chronic nightmares. For example, as one woman gained mastery over her snake phobia, she dreamt that the boa constrictor befriended her and was helping her to wash the dishes! Reptiles soon faded from her dreams. The changes endured.

The 1960s ushered in remarkable transformative changes in the explanation and modification of human functioning and change (Bandura, 2004b). Causal analysis shifted from unconscious psychic dynamics to transactional psychosocial dynamics. Human functioning was construed as the product of the dynamic interplay between personal, behavioral, and environmental influences. Action-oriented treatments replaced interpretive talk therapies. The modes of treatment were altered in the content, locus, and agents of change. Within a decade, the field was transformed by a major paradigm shift. New conceptual models and analytic methodologies were created. New sets of periodicals were launched for the rising stream of interest. New organizations were formed for the advancement of behaviorally oriented approaches. New professional conventions provided a forum for the exchange of ideas.

Not all the critics of the psychodynamic model worshipped at the same theoretical alter, however. Some thought the operant conditioning route provided the best glimpse of the promised land. Others adopted Hullian theory. I took the social cognitive route, emphasizing the influential role of agentic capabilities in self-development, adaptation,

and change. Vigorous battles were fought over cognitive determinants and their scientific legitimacy (Bandura, 1995b, 1996; Catania, 1975; Skinner, 1977).

The popular media were deluging the public with repugnant imagery of brainwashing and frightful scenarios of *1984* and *Brave New World* dominated by social engineers wielding powerful methods of behavioral control. The hit movie, *A Clockwork Orange*, graphically portrayed the fiendish nature of behavior modifiers physically shocking people into submission. In his movie *Sleeper*, Woody Allen amusingly outwits the ironclad control by despotic social engineers who reduce humans to mindless zombies. Skinner's (1971) publication, *Beyond Freedom and Dignity*, alarmed the public that the application of these new psychological methods would strip people of their dignity and deprive them of their freedom. The Unabomber targeted Jim McConnell at the University of Michigan as his first victim with a tirade about the evils of behavior modification. Lyndon La Rouche, who became a perennial candidate for the US presidency, branded the practitioners of behavioral approaches as "Rockefeller Nazis," formally tried some of the leading figures in his tribunal for crimes against humanity, stormed classes at the University of New York at Stony Brook, and issued threats requiring police surveillance of the Associates for the Advancement and Behavior Therapy (AABT) convention in Chicago. As in any professional practice, there were some reprehensible applications of behavioral principles, especially in coercive institutional systems, that affirmed and fueled the public's fears.

At the height of this media frenzy, I began my term as president of the American Psychological Association (APA). A responsible social science must concern itself not only with the advancement of knowledge but also with the effects of its social applications. In keeping with this dual commitment, we formed an APA interdisciplinary task force to examine the way in which knowledge on behavioral modification was being used both at the individual and institutional level. Its wide-ranging analysis, which was published in the volume, *Ethical Issues in Behavior Modification* (Stoltz, 1978), provided a thoughtful evaluation of existing applications and a set of standards for ethical practice that helped to dispel the frightful misconceptions propagated by the mass media. Growing applications of [p. 357 ↓] cognitive behavioral treatments not only won public acceptance, but are now used as the evidence-based method of choice for diverse maladies of the human condition. This fascinating Odyssey involved dual

transformative changes – a paradigm shift in theory and practice as well as a sweeping change in public acceptance.

Self-Efficacy Component in Social Cognitive Theory

My entry into self-efficacy was serendipitous. In the development and evaluation of the guided mastery treatment, we focused on three fundamental processes: the power of the treatment to promote psychosocial changes; the generality or scope of the effected changes; and their durability or maintenance. Having demonstrated the power of this mode of treatment on each of these evaluative dimensions, I explored the possibility of a further function, the power of a treatment to build resilience to adverse experiences. We tested the proposition that a lot of neutral or positive experiences after functioning is fully restored can neutralize the negative impact of an aversive event and curtail the spread of negative effects.

In a follow-up assessment, the participants expressed deep gratitude to be rid of their phobia, but then explained that the treatment had a more profound impact. Their lives had been debilitated intrapsychically, socially, recreationally, and occupationally for 20 to 30 years. They were plagued by recurrent nightmares and perturbing ruminations. To overcome, within a few hours, a phobic dread that had constricted and tormented their lives was a transformational experience that radically altered their beliefs in their efficacy to exercise control over their lives. They were acting on their new self-efficacy belief and enjoying their successes, much to their surprise. These preliminary findings pointed to a common mechanism through which personal agency is exercised.

I mounted a multifaceted program of research to gain a deeper understanding of the nature and function of this belief system. To guide this new mission, the theory addressed the key aspects of perceived self-efficacy (Bandura, 1997). These include the origins of efficacy beliefs, their structure and function, their diverse effects, the processes through which they work, and the modes of influence by which a resilient sense of efficacy can be created and strengthened for personal and social change. Diverse lines of research, conducted by a variety of investigators, provided new insights

into the role of perceived self-efficacy in the fields of education, health promotion and disease prevention, clinical dysfunctions such as anxiety disorders, depression, eating disorders, substance abuse, personal and team athletic attainments, organizational functioning, and the efficacy of our social and political systems to make a difference in people's lives.

A major question in any theory of cognitive regulation of motivation, affect, and action concerns the issue of causality. A variety of experimental strategies were used to verify that beliefs of personal efficacy are contributors to performance, not merely reflectors of it (Bandura, 1997; Bandura and Locke, 2003). The issues raised by devout proponents of competing theories were also addressed both conceptually and empirically (Bandura, 2008a).

I receive a steady flow of e-mails requesting my all-purpose measure of self-efficacy or a couple of trait-like items that could be inserted into an omnibus questionnaire. There is no all-purpose measure of self-efficacy. The requesters are referred to a detailed instructional manual on how to construct psychometrically sound self-efficacy scales (Bandura, 2006d). Self-efficacy assessments are tailored to spheres of functioning and the realities people have to manage. Another entry in the research agenda was to differentiate an agentic model of personality from a trait model. It also required purging misconceptions of constructs. *Self-efficacy* is a judgment of personal capability; *self-esteem* is a [p. 358 ↓] judgment of self-worth. These are entirely different constructs. Nor is self-efficacy the same as *locus of control*, which is a belief about whether outcomes are contingent on behavior or on extraneous factors. Belief that outcomes are determined by one's performance (internal locus) can be motivating under high self-efficacy but demoralizing under low self-efficacy to produce the required performance.

The field of personality is deeply entrenched in trait thinking. The currently-in-vogue "Big-Five" model shrinks personality to five clusters of behaviors disconnected from the vast body of knowledge on the development, organization, regulation, and modification of behavior (McCrae and Costa, 1996). These traits are measured by decon-textualized behavioral descriptors in a one-size-fits-all questionnaire. Human functioning is too multifaceted, contextualized, and conditionally manifested to be reduced to a small number of behavioral descriptors reified as personality determinants.

A five-fold behavioral taxonomy is hardly a theory of personality. The traits comprising this approach should be renamed behavioral traits rather than personality traits because the items are mainly clusters of behaviors. One needs a theory of personality to explain how intrapersonal factors contribute to the development and adoption of conscientious, agreeable, receptive, and socially outgoing behavior. In a chapter entitled "Social Cognitive Theory of Personality," I argued that personality determinants reside in intrapersonal factors and psychosocial processes not in behavioral clusters (Bandura, 1999b). The convenience of all-purpose global tests of personal attributes is gained at the cost of explanatory and predictive power. All too often, personality psychology is marginalized as simply a supplier of handy off-the-shelf trait measures. Such measures are being appended, often with little conceptual rationale, to whatever one is studying under the illusion that it represents the contribution of "personality" to human functioning.

Evolution of an Agentic Theory of Human Behavior

There are four core properties of human agency (Bandura, 2006a, 2008a). One such property is *intentionality*. People form intentions that include action plans and strategies for realizing them. The second property involves the temporal extension of agency through *forethought*. This includes more than future-directed plans. People set themselves goals and anticipate likely outcomes of prospective actions to guide and motivate their efforts anticipatorily. When projected over a long time course, a forethoughtful perspective provides direction, coherence, and meaning to one's life. The third feature is *self-reactiveness*. Agents are not only planners and forethinkers. They are also self-regulators. They adopt personal standards and monitor and regulate their actions by evaluative self-reactions. They do things that give them satisfaction and a sense of self-worth, and refrain from actions that bring self-censure. The fourth feature is *self-reflectiveness*. People are not only agents of action. They are self-examiners of their own functioning. Through functional self-awareness they reflect on their personal efficacy, the soundness of their thoughts and actions, the meaning of their pursuits, and try to make corrective adjustments if necessary.

Locus of Causation

Over the years, theorists engaged in vigorous debates on whether the causes of human behavior reside in the environment as the situationists claim, or in the individual as the dispositionalists claim. Although most theorists now adopt an interactionist model of causation, we are still arguing over whether the person or the situation is the locus of causation. Sociologists and many social psychologists argue for the power of the situation, while personologists argue for the power of the person. In the interactionist view, [p. 359 ↓] human functioning is a product of the interaction between personal and environmental influences. However, there are three types of interactionism, two of which subscribe to one-way causation in the link to behavior.

In the *unidirectional* model, persons and situations are treated as independent influences that combine in unspecified ways to produce behavior. The major weakness with this causal model is that personal and environmental influences do not function as independent determinants. They affect each other. People create, alter, and destroy environments. The changes they produce in environmental conditions, in turn, affect them personally. The causally unidirectional relation to behavior is another serious deficiency of this model of interactionism.

The *partially bidirectional* model of interaction, which is widely adopted in personality theory, acknowledges that persons and situations affect each other, but still treats influences relating to behavior as flowing in only one direction. The person-situation interchange is said to produce behavior unidirectionally, but the behavior itself does not affect the ongoing transaction between the person and the situation. A major limitation of this causal model is that behavior is not procreated by an intimate interchange between a behaviorless person and the environment. Such a feat would be analogous to immaculate conception. Except through their social stimulus value, people cannot affect their environment, other than by their actions. Behavior is an interacting determinant rather than a detached byproduct of a behaviorless person-situation interchange.

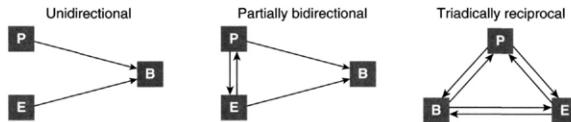
The partially bidirectional model of interactionism is typically evaluated by partitioning the average amount of variance in behavior attributed to persons, environments,

and their interaction. In this conception, the interactional form is statistical not an interdependently dynamic one in the causal structure. The statistical partitioning reminds one of the nonswimming statistician who drowned while crossing a river that averaged two feet in depth.

Social cognitive theory conceptualizes the interactional causal structure as *triadic reciprocal causation*. In this conception, human functioning is a product of a reciprocal interplay of intrapersonal, behavioral, and environmental determinants (Bandura, 1978). Phillips and Orton (1983) claimed that I was proposing a theory of holistic causation in which the three major classes of determinants act simultaneously as a fused whole. They likened the triadic reciprocity to Haldane's holistic doctrine that William James (1884) criticized. This characterization came as a surprise to me because at no time have I advocated a doctrine of simultaneous holistic reciprocity in which the triadic determinants operate "at the same moment in time."

In a rejoinder, I explained that a causal process in which the interactants are influencers and influenced at the same time is not only illogical but functionally impossible (Bandura, 1983). A causal factor takes some time to exert influence and experience a reciprocal effect. In a verbal interchange, for example, questions and answers cannot occur simultaneously. Mutual influences and their reciprocal effects do not spring forth all at once. They work their mutual effects over [p. 360 ↓] variable time courses. Time lags between events permit clarification of how different segments of reciprocity operate. Knowledge of how the various subsystems work advances understanding of how the superordinate system operates. It is the subsystems and their various interactions rather than the entirety that are analyzed. In the analytic decomposition of triadic causation, different subspecialties of psychology center their inquiry on selected segments of the reciprocal interplay (Bandura, 1986). A herculean effort to examine every possible interactant at the same time would beget investigatory paralysis.

Figure 17.1 Schematization of different causal models of interaction between key classes of determinants. Social cognitive theory is founded on the causal model of triadic reciprocal determination



Staddon (1984), a proponent of the operant conditioning view, got in on the act as well. He argued that internal determinants are unmeasurable so causal structures should be confined to stimulus inputs. Scientific advances are promoted by two kinds of theories (Nagel, 1961). One form seeks relations between directly observable events but shies away from the mechanisms subserving the observable events. The second form focuses on the mechanisms that explain the functional relations between observable events. The problem with the orthodox version of environmental determinism is that behavior is neither always cued by the stimuli that precede it nor always controlled by the stimuli that follow it. Hence, Staddon expanded this model to one in which behavior is under dual stimulus control – by current external stimuli and the internal residuum of past stimulus inputs. The implanted history of reinforcement carried the major explanatory burden. In a commentary, I explained that an implanted history is an inferred inner determinant, not a directly observable one (Bandura, 1984). The dispute is no longer about inner causes but the form they take. I documented the diverse ways in which cognition breaks the chain of stimulus control. Forethought can enhance, neutralize, or override the impact of environmental inputs.

There is much advice on how to conduct informative research and get it published, but little on how to manage challenges to, and misunderstandings of, posited theories and the findings of verification tests. This is part and parcel of a scientific career that can command a lot of one's attention. If the theory addresses diverse aspects of human functioning and is widely cited, the disputational aspect of professional life can crowd a busy academic schedule. I discuss elsewhere issues regarding the process of theory building (Bandura, 2005a).

Human Capacity for Self-Regulation

As previously noted, people's capacity to regulate their own functioning and shape the course their lives take is a core feature of social cognitive theory. The road I

have traveled is very much in keeping with the agentic perspective toward human development, adaptation, and change. It underpins social cognitive theory (Bandura, 2006c). I grew up in a tiny rural hamlet in northern Alberta. The only school in town, which housed first grade through high school, was woefully short of teachers and educational resources. Two teachers had to teach the entire high school curriculum. They tried their best but were not always fully informed in key subject areas. We had to take charge of our own learning. Self-directed learning was an essential means of academic self-development, not a theoretical abstraction debated in arcane language in learned circles. However, the paucity of educational resources turned out to be enabling rather than handicapping. The content of courses is perishable, but self-regulatory skills have lasting functional value whatever the pursuit might be. These formative experiences spoke to an agentic view of human behavior that accented positive enabling factors rather than debilitating risk factors (Bandura, 2006a, 2008b).

In our excursion into the nature of self-directedness, our laboratory studies explored the mechanisms of self-regulation (Bandura, [p. 361 ↓] 1986, 1991c). Some of the studies clarified how personal standards are constructed from the myriad social influences. Others documented the regulatory power of self-reactive influences. Rational models of human behavior embraced the centrality of agency but they too provided a truncated view of self-regulation rooted in the market metaphor. Behavior was said to be regulated by self-interest construed almost entirely in terms of material costs and benefits. We demonstrated that human motivation and performance attainments are governed not only by material incentives, but also by social incentives, and self-evaluative incentives linked to personal standards. People often settle for alternatives of marginal utility or even sacrifice material gain to preserve their positive self-regard. Some of our studies examined self-regulation under conflictual conditions where individuals are rewarded for behavior they devalue, or are punished for activities they personally value. Principled dissenters often find themselves in the latter predicament. Their sense of self-worth is so strongly invested in certain convictions that they will submit to maltreatment rather than accede to what they regard as unjust or immoral.

Operant conditioners defined self-regulation out of existence by rechristening it as “stimulus control” and locating it in the external environment. They took me to task for introducing self-referent factors into the determination of behavior (Catania, 1975). In rejoinders I relocated self-management in the sentient person and reviewed the

growing body of evidence on the means by which individuals exercise self-directedness (Bandura, 1976).

This was not a hospitable time to present an agentic theory of human behavior. Psychodynamicists depicted behavior as driven unconsciously by impulses and complexes. Behaviorists portrayed behavior as shaped and shepherded by environmental forces. The cognitive revolution was ushered in on a computer metaphor. This conception stripped humans of agentic capabilities, a functional consciousness, and a self-identity. The mind as a symbol manipulator in the likeness of a linear computer became the conceptual model for the times. It was not individuals but their subpersonal parts that were orchestrating activities nonconsciously. Control theories of motivation and self-regulation (Carver and Scheier, 1981; Lord and Levy, 1994) drew heavily on Powers' (1973) perceptual control theory, which is an outgrowth of the mechanical cybernetic model. In addition, as will be shown later, eliminative materialists likened cognitive factors, such as beliefs, goals, and expectations, to the phlogiston of yesteryear.

A good share of the conceptual sparring over the nature of self-regulation involved differences between social cognitive theory and Powers' control theory. The core feature of Powers' theory is the negative feedback loop. Discrepancies between a programmed reference standard and the perceived input from the output automatically trigger action to match the standard. Error correction is the driving force. I regarded self-regulation by negative discrepancy as only half the story and, in many respects, the less challenging half to explain.

Viewed from the agentic perspective, self-regulation operates through dual control systems – proactive discrepancy production working in concert with reactive discrepancy reduction (Bandura, 1991c). People are aspiring and proactive organisms, not just reactive ones. Their capacity to exercise forethought enables them to wield adoptive control anticipatorily rather than being simply reactive to the effects of their efforts. We demonstrated that people motivate and guide their actions through proactive control by setting themselves challenging goals and performance standards that create negative discrepancies to be mastered. They then mobilize their effort and resources to fulfill those challenges. After people attain the goals they have been pursuing, those of high perceived self-efficacy set a higher standard for themselves. The adoption

of further challenges creates new motivating discrepancies to be mastered. In short, people are motivated and [p. 362 ↓] guided by foresight of goals, not just hindsight of shortfalls.

A theory of self-regulation governed by forethought and affective self-reactions did not sit well with Powers (1991), the foremost advocate of control theory. In his view, the human organism is “nothing more than a connection between one set of physical quantities in the environment (input quantities) and another set of physical quantities in the environment (output quantities)” (Powers, 1978: 421). Cognitive and affective processes were dismissed as irrelevant because “we are not modeling the interior of the subject” (1978: 432). We questioned the conceptual and empirical adequacy of this austere mechanistic model, as well as derivative control theories that grafted on the negative feedback loop a variety of intrapersonal properties (Bandura, 1991b; Bandura and Locke, 2003; Locke, 1994). These adjuncts violated the parent control theory but befitted a sentient being.

Social cognitive theory lends itself readily to social applications. Collaboration with Robert De Busk and Kate Lorig in the Stanford Medical School provided an opportunity to extend the agentic model of self-regulation to health promotion and disease prevention. These applications were rooted in a shift of emphasis from a disease model to a health model. It promotes effective self-management of health habits by cultivating self-regulatory skills that enable people to live healthier and retard the process of aging. This work led to the development of a self-management health system that promotes health and reduces risk of disease on a large scale at relatively low cost (Bandura, 2004c, 2005b).

Vast populations have no access to services that promote health and timely help in changing habits that impair it. By linking the interactive aspects of the self-management model to the Internet, one can vastly expand its availability to people wherever they may live. The goal of this implementational extension is to develop interactive online systems that enable people worldwide to improve the quality of their health.

Self-Regulation in the Exercise of Moral Agency

In areas of functioning involving achievement strivings and cultivation of competencies, the personal standards that serve as the mark of adequacy are progressively altered as knowledge and skills are acquired and challenges are met. In many areas of social and moral conduct, the internal standards are relatively stable. That is, people do not change from week to week in what they regard as right or wrong or good or bad. The agentic theory of self-regulation encompasses not only aspirational self-management but moral self-regulation as well in the exercise of moral agency (Bandura, 1991, 2004a). The self-regulation operates through the same basic set of subfunctions across aspirational and moral spheres of functioning. The verified explanatory commonality of the self-regulatory theory across markedly diverse spheres of activities is in keeping with Occam's maxim advocating theoretical parsimony.

Psychological theories of morality focus heavily on the acquisition of moral standards and the structure of abstract moral reasoning, often to the neglect of moral conduct. A complete theory of moral agency must link moral knowledge and reasoning to moral action. This requires an agentic theory of morality rather than one confined mainly to cognitions about morality. Adoption of moral standards does not create an immutable internal moral control system. Indeed, large-scale humanities are typically perpetrated by people who are considerate and compassionate in other aspects of their lives. A full theory of moral agency must, therefore, explain how otherwise considerate people can behave inhumanely.

In the social cognitive theory of moral agency, there are eight psychosocial mechanisms operating at four loci in which moral self-sanctions can be selectively disengaged from harmful practices (Bandura, 1999a). At the *behavior locus*, worthy ends are used to sanctify harmful means by social and moral justification, exonerative comparison that [p. 363 ↓] renders the practices righteous, and sanitizing and convoluted language that disguise what is being done. At the *agency locus*, people obscure personal responsibility by displacement and diffusion of responsibility. This absolves them of accountability for the harm they cause. At the *outcomes locus*, the detrimental

social effects of one's actions are ignored, minimized, or disputed. At the *victim locus*, perpetrators dehumanize and blame recipients for bringing the maltreatment on themselves. These mechanisms operate, often in concert, at both individual and social system levels.

The philosopher Seneca once portrayed seeming serendipity as: "Luck is what happens when preparation meets opportunity." This account characterizes well my varied partnerships in research designed to advance understanding of the determinants, mechanisms, and effects of moral disengagement in diverse spheres of life. One such collaborative effort is a multifaceted longitudinal project with Gian Caprara and his associates, Claudio Barbaranelli and Tina Pastorelli at the University of Rome "La Sapienza," in which we are studying child development from an agentic theoretical perspective. Part of this research centered on the development and exercise of moral agency. We created scales for measuring facility in moral disengagement and demonstrated that moral disengagement increases proclivity for aggressive and antisocial activities.

A San Francisco newspaper ran a story on the program of research by Lisa Bero, at the University of California Medical School, showing that corporate funding of research biases the findings. I checked with Lisa on whether she would be interested in a large-scale analysis of moral disengagement in the corporate world. A joint project we mounted verified the widespread moral disengagement at the social systems level in the tobacco, lead, vinyl chloride, and silicosis-producing industries. Through Caprara's acquaintance with Lazlo Zsolni, at the Business Ethics Center at Budapest University, we extended the analysis of corporate moral disengagement to the massive Bhopal chemical disaster, Nestlé's aggressive marketing of infant formula products to third-world countries despite serious health effects, the defective Ford Pinto that took a heavy toll of lives, and the linguistic sanitizing of the mishap at the nuclear power plant on Three Mile Island.

Michael Osofsky, one of my undergraduate advisees at Stanford, often accompanied his father who chaired the Psychiatry Departments at Louisiana State University, on consulting visits to a Southern maximum security penitentiary. Because of this relationship, the warden not only granted us permission to study members of the execution team but also gained us access to two other Southern penitentiaries.

This research gave us a better understanding of how executioners disengage moral self-sanctions from the taking of human life. Haney (1997) documented the morally disengaging ways in which capital trials are structured and conducted to enable jurors to sentence a person to death. A collaborative public survey study with Alfred McAlister at the University of Texas testified to how moral disengagement eases the public's qualms about the use of state executions. This set of studies furthered understanding of how moral disengagement is enlisted at each of the three levels in the application of the death penalty – at the societal, judicial, and execution levels.

In another line of study in our research partnership, McAlister and I were conducting a national survey on moral disengagement in support of military force. Midway through this study the terrorists struck the World Trade Center and the Pentagon. The terrorist attacks raised the level of moral disengagement. In path analysis, moral disengagement completely mediated the impact of the terrorist attack and sociodemographic characteristics on public support for the use of military force. These diverse collaborative projects illustrate the multiplicative advances in knowledge that can be achieved by courting opportunities when they arise. I also analyzed independently moral disengagement in the commercialization of gratuitous violence by the broadcast [p. 364 ↓] industry (Bandura, 1973), in terrorism and in the enlistment of public support for, and conduct of, the war in Iraq (Bandura, 2004d).

My laboratory research on the moral disinhibitory function of dehumanization and diffusion of responsibility predated the studies cited above of moral disengagement at the social systems level (Bandura et al., 1975). While demonstrating how easy it is to bring out the worst in others, these experiments also revealed the power of humanization to curb inhumane conduct. However, I discontinued the laboratory studies on the disengagement of moral agency because it required the experimenters to disengage their own moral qualms to be able to them.

Affirmation of Agency in a Reductionistic Era

There are two theoretical developments that minimize or dismiss the exercise of human agency. The first detractor is physical eliminationism. In this view, human behavior is governed by intricate neural networks that operate outside one's awareness and control. The second is the growing geneticization of human behavior. Social roles and human practices are increasingly being proclaimed as driven by prehistoric biological programming. In both of these theoretical orientations, human behavior is shaped and regulated at the subpersonal level. I addressed elsewhere, at some length, these two lines of theorizing and critically analyzed some of the widely cited evidence for them (Bandura, 2008a). A brief summary of this analysis is presented in the sections that follow.

Physicalistic Theory of Human Agency

Agentic contributions to human functioning are dismissed in some quarters on the grounds that human behavior is regulated by neuronal mechanisms operating at a subpersonal level. Thoughts are construed as epiphenomenal events that create an illusion of control but actually have no effect on how one behaves. In this view, humans are essentially conscious hosts of automata that dictate their behavior subpersonally. In support of an agentic theory of human functioning, I argued that physical eliminationists frame the issue of personal regulation in the wrong terms at the wrong level of control.

In acting as agents, individuals obviously are neither aware of, nor directly control, their neuronal mechanisms. Rather, they exercise *second-order control*. They do so by intentionally engaging in activities at the macrobehavioral level known to be functionally related to given outcomes. In pursuing these activities, over which they can exercise control, they shape the functional circuitry and enlist the neurophysiological events subserving their pursuits. Cognitions are higher-level cerebral events involving deliberative, reflective, self-referential, and evaluative processes operating in a top-

down fashion through highly interconnected brain systems within the same material entity rather than in a physicalistic dualism.

Consider the following analogy. In driving an automobile to a desired place, the driver engages in coordinated acts of shifting gears, steering, manipulating the gas pedal, and applying brakes. These deliberate acts, which the driver controls directly, regulate the mechanical machinery to get safely to where the driver wants to go. But the driver has neither awareness nor understanding of the correlative microcombustion, transmission, steering, and braking processes subserving the driver's purposes. The deliberate planning of where to go on a trip, what route to take, what to do when one gets there, and securing reservations for these diverse activities far in advance requires considerable proactive top-down cognitive regulation. The temporal structuring of behavior by goals and purposes sets the course for one's activities. Proximal self-regulation provides the guides, strategies, and motivators in the here and now to get to where one is going [p. 365 ↓] (Bandura, 1991c). Having constructed a trip, travelers cannot sit back and wait for lower-level sensory-motor activity to consummate the arrangements unconsciously.

Similarly, in second-order control over their cardiovascular functioning, individuals obviously do not intentionally direct their atrial and ventricular cardiac muscle fibers to fire and their aortic and pulmonary valves to open and close. However, by intentionally engaging in an exercise routine and controlling their activity level, they can enhance their cardiac functioning and raise and lower their heart rate without having the foggiest idea of how they indirectly recruited the subserving neurophysiological mechanisms. In short, enactments of functional activities at the controllable macrobehavioral level serve as the means for agentic recruitment of the subserving events at the microneural level.

Almost everyone adopts the ontological view that cognitive events are brain activities not immaterial entities. It is the epistemological form of reductability that is most in contention. The major argument against it is that each level of complexity – atomic, molecular, biological, psychological, and social structural – involves emergent new properties that are distinct to that level and, therefore, must be explained in their own right; for example, knowing the locality and brain circuitry subserving learning can say little about how best to devise conditions of learning in terms of level of abstractness and challenge; how to provide incentives to get people to attend to,

process, and organize relevant information; and whether learning is better achieved independently, cooperatively, or competitively. The optimal conditions must be specified by psychological principles. There is little at the subatomic or neuronal level that can tell us how to develop efficacious parents, teachers, and social reformers or how to build and run social systems.

People are contributors to their activities, not just onlooking hosts of subpersonal networks autonomously shaping and regulating their performances. An aspiring violinist, for example, has to practice tenaciously to train the brain, build muscular strength and dexterity, and hone sensory acuity to realize a virtuoso performance. Tell an aspiring violinist, who has spent countless hours training the brain and manual fingering and bowing dexterity to execute the pyrotechnical wizardry of a Paganini violin concerto, that the neural network is really the violinist and the indefatigable musician is just a self-aggrandizing illusionist.

The sensory, motor, and cerebral systems are tools people use to accomplish the tasks and goals that give meaning, direction, and satisfaction to their lives. To make their way successfully through a complex world people have to make sound judgments about their capabilities, anticipate the probable effects of different events and courses of action, size up sociostructural opportunities and constraints, and regulate their behavior accordingly. These belief systems are a working model of the world that enables people to achieve desired futures and avoid untoward ones.

There is growing unease about progressive divestiture of different aspects of psychology to biology and subpersonal neuroscience. It is feared that as we give away more and more psychology to disciplines lower down the conceptual food chain, there will be no core psychological discipline left. Psychology will become merely a branch of biology. Contrary to the proclamations of the divestitive oracles, psychology is the one discipline that uniquely encompasses the complex interplay among intrapersonal, biological, interpersonal, and sociostructural determinants of human functioning. Psychology is best suited to advance understanding of the integrated biopsychosocial nature of humans, and how they manage and shape the everyday world around them. The field of psychology should be articulating a broad vision of human beings, not a reductive fragmentary one.

Growing Primacy of Human Agency in the Co-Evolution Process

The conceptions of human nature regarding the capacity to exercise some measure of [p. 366 ↓] control have changed markedly over time. In the early theological conceptions, human nature was ordained by original divine design. Evolutionism transformed the conception to one in which human nature is shaped by environmental pressures acting on random gene mutations and reproductive recombinations. This process of natural selection is devoid of deliberate plans or purposes. The evolutionary emergence of language and abstract and regulative cognitive capacities transformed the nature of the coevolution process. It provided the neuronal structure for supplanting aimless environmental selection with cognitive agency. Human forebears evolved into a sentient agentic species. Their advanced symbolizing capacity enabled humans to transcend the dictates of their immediate environment and made them unique in their power to shape their circumstances and life. Although not limitless, learnability and agentic capability became the hallmark of human nature.

Biology provides the information-processing architectures and potentialities and sets constraints. But in most spheres of functioning, biology permits a broad range of cultural possibilities. As Jay Gould (1987) notes, the major explanatory battle is not between nature and nurture as commonly framed. But whether nature operates as a determinist, that has culture on a “tight leash,” as Wilson (1998) contends, or as a potentialist that has culture on a “loose leash,” as Gould (1987) maintains. Biological determinists support a conservative view of society. It emphasizes the rule of nature, inherent constraints, and limitations. They argue that people should not try to remake themselves and their societies against the rule of nature, however they construe it. Biological potentialists give greater weight to enabling social conditions for personal development and societal change.

Evidence supports the potentialist view. Humans have created societies of diverse natures: aggressive and pacific ones, egalitarian and despotic ones, altruistic and selfish ones, individualistic and collectivistic ones, enlightened and backward ones. The human species exhibits transformative changeability as well as interculture and

intraculture diversity. People have changed little genetically over the past millennium but they have changed markedly even over the recent decades in their beliefs, mores, social and occupational roles, cohabiting arrangements, family practices, and styles of behavior. They have done so through rapid cultural and technological evolution. Cultures evolve over generations and shape the ways people need to live in the particular cultural milieu in which they are immersed.

Other species are heavily innately programmed for stereotypic survival in a particular habitat. In contrast, as an agentic species, humans devise ways of adapting flexibly to remarkably diverse geographic, climatic, and social environments. Consider the many ways in which the psychosocial side of coevolution is gaining ascendancy through agentic ingenuity. People create technologies to transcend their biological limitations. For example, humans have not evolved morphologically to fly but they are soaring through the air and even in the rarified atmosphere of outer space at breakneck speeds, despite the biological unachievability. Agentic inventiveness overrode biological design in getting them airborne. People also use their ingenuity to circumvent and insulate themselves from selection pressures. They create devices that compensate immensely for their sensory and physical limitations. They transcend time, place, and distance as they interact globally with the virtual environment of the cyberworld.

Through genetic engineering, humans are creating biological natures for better or for worse, rather than waiting for the slow process of natural evolution. They are now changing the genetic make-up of plants and animals that evolved over eons. Not only are humans cutting and splicing nature's genetic material, but, through synthetic biology, they are also creating new types of genomes. In short, humans are an agentic species that is altering evolutionary heritages and shaping their future.

[p. 367 ↓] The notion in vogue is that biological evolution provides the potential and culture can do only so much within those constraints. This view flies in the face of the growing primacy of human agency in the coevolution process. It is not that social cognitive theory minimizes biological endowment. Quite the contrary. People have evolved the complex biological systems required for the very agentic activities by which personal and social changes are realized. These include deliberative and generative thought, fore-thoughtful self-regulation, and reflective self-evaluation. Neither the agentic human ascendancy in the coevolution process nor the rapid transformational

societal changes would be possible without the biological endowment of abstract cognitive capabilities.

Social cognitive theory highlights the forward-looking impact of our biological endowment, rather than backward-looking speculation about adaptation to primitive conditions of prehistoric times. The study of how humans are changing endowed heritages, circumventing biological constraints, and shaping their future through social and technological evolution is more fruitful than spinning fanciful stories about prehistoric mating patterns in drafty caves.

Were Darwin writing today, he would be documenting the overwhelming human domination of the environment. As the unrivaled ruling species atop the food chain, we are degrading the earth's life support systems and drafting the requiem for biodiversity. By wielding powerful technologies that amplify control over the environment, driven by a foreshortened perspective, humans may be well on the road to outsmarting themselves into irreversible ecological crises that they can no longer control.

Exercise of Agency in Cultural Context

Contentious dualisms pervade our field, pitting autonomy against interdependence, individualism against collectivism, and social structure against agency. Cultures are dynamic and internally diverse systems, not static monoliths (Bandura, 2002b, Kim et al., 1994). Analyses across activity domains and classes of social relationships reveal that people behave communally in some aspects of their lives and individualistically in many other aspects. They express their cultural orientations conditionally depending on incentive conditions, rather than invariantly. The categorical approach masks extensive diversity.

Not only are cultures not monolithic entities but they are no longer insular. Transnational interdependencies and global market forces are restructuring national economies and shaping the political and social life of societies. Advanced telecommunications technologies are disseminating ideas, values, and styles of behavior globally at an unprecedented rate. The symbolic environment, feeding off communication satellites, is altering national cultures and homogenizing collective consciousness. People are

spending much of their time in the ever-expanding cyberworld. This is furthering the globalization of culture. In addition, mass migrations of people and high global mobility of entertainers, athletes, journalists, academics, and employees of multinational corporations are changing cultural landscapes. This intermixing creates new hybrid cultural forms, blending elements from different ethnicities. These social forces are homogenizing some aspects of life, polarizing other aspects, and fostering a lot of cultural hybridization.

It is widely claimed that Western theories lack generalizability to non-Western cultures (Sampson, 1977). One must distinguish between basic human capacities and how culture shapes these potentialities into diverse forms. For example, social modeling through observational learning figures prominently in social cognitive theory. Humans have evolved an advanced capacity for learning through observation of modeled attitudes, values, and styles of behavior. It is essential for their [p. 368 ↓] self-development and functioning regardless of the culture in which they reside. Modeling is a universalized human capacity. But what is modeled, how modeling influences are structured, and the purposes they serve, varies in different cultural milieus.

The same distinction in levels of analysis applies to perceived efficacy. A common duality inappropriately equates self-efficacy with self-centered individualism and pits it against collectivism (Schooler, 1990). Because efficacy beliefs involve self-referent processes, self-efficacy is often misconstrued as aggrandizement of an individualistic autonomous self and contrasted with an interdependent collectivistic self. Self-efficacy does not come in only an individualistic form nor with a built-in value system. People's belief in their efficacy is exercised in individual, proxy, and collective forms. Social cognitive theory is, therefore, just as relevant to human attainments realized through interdependent collective effort as to those achieved individually. Moreover, self-efficacy can serve communal purposes as well as individualistic ones.

Being immobilized by self-doubt about one's capabilities and belief in the futility of effort has little adaptive advantage. A growing body of research shows that, indeed, a resilient sense of efficacy has generalized functional value regardless of whether one resides in an individualistically oriented culture or a collectivistically oriented one (Bandura, 1995a, 2002b). But how efficacy beliefs are developed, the form they take, the ways in which they are exercised, and the purposes to which they are put vary cross-culturally.

The cross-cultural findings debunk the misconception that belief in one's efficacy is an egocentric orientation wedded to Western individualism. Personal efficacy is valued, not because of reverence for individualism, but because a resilient sense of efficacy has generalized functional value regardless of whether activities are pursued individually or by people working together for common cause.

In short, there is a cultural commonality in basic agentic capacities and mechanisms of operation, but diversity in the culturing of these inherent capacities. In this dual-level analysis, universality is not incompatible with manifest cultural plurality. Cultural variations emerge from universalized capacities through the influence of social practices reflecting shared values and norms, incentive systems, role prescriptions, and pervasive modeling of distinctive styles of thinking and behaving. Kluckholn and Murray summarized eloquently the blend of universality, commonality, and uniqueness of human qualities (Muñoz and Mendelson, 2005): "Every person is in certain aspects like all other people. Like some other people. Like no other person."

Going Global with Social Cognitive Theory

Global applications of social cognitive theory to promote society-wide change testify to the effectiveness of social modeling and enhanced individual and collective efficacy to improve the quality of life in diverse cultural milieus (Bandura, 2006b, 2006c).

These applications, which reach millions of people in Africa, Asia, and Latin America, address some of the most urgent global problems. Soaring population growth tops the list because of its heavy strain on natural and social resources and on ecological supports of life. Developed nations are stabilizing their population, but developing ones, where most of the growth is occurring, are rapidly doubling their populations. We are currently at 6.7 billion, adding about one billion every fifteen years, and heading toward a population of nine to ten billion by mid-century. We are witnessing spiraling destruction of ecological systems that have evolved over eons that keep our planet cool and habitable. Burgeoning population growth also fuels civil strife. Most violent conflicts are in countries with large youth populations that are uneducated, unemployed, and living in poverty under autocratic rulers plagued by corruption. Youth living [p. 369 ↓] under such dismal life conditions are easily recruited for violent activities.

Another widespread problem is the pernicious gender inequality. In these societies, women are marginalized, devalued, disallowed aspiration, and denied their liberty and dignity. The spreading AIDS epidemic is another mounting global problem with devastating societal consequences. Some societies present unique problems that require special social themes tailored to the detrimental cultural practices.

Long-running serialized dramas serve as the vehicle for promoting personal and social changes. To change deeply held beliefs and social practices requires strong emotional bonding to enabling models that exemplify a vision of a better future and realistic paths to it. There are no quick fixes. The dramatic productions are not just fleeting fanciful stories. They dramatize people's everyday struggles, the impediments they face, and the effects of different societal practices. The storylines speak to people's fears and hopes, help them see a better life, and inform and enable them to take the steps to realize it.

The storylines model family planning, women's equality, degrading dowry systems, spouse abuse, environmental conservation, AIDS prevention, and varied life skills. Hundreds of episodes get people emotionally engaged in the evolving lives of the models and are inspired and enabled by them. This is a highly flexible format, which contributes to its generalizability, versatility, and power. By including multiple intersecting plotlines, one can address diverse aspects of life at both the individual and social structural level.

Major advances in science and its applications require the combination of the expertise of diverse sources. This is best achieved, not through bureaucratic interdisciplinary prescription, but through collaboration fueled by common interest and purpose. These ambitious global applications testify to the value of informally created partnerships. Promoting large-scale changes requires three component models – theoretical, translational, and social diffusion models. The first component is a *theoretical model*. It specifies the determinants of psychosocial change and the mechanisms through which they produce their effects. This knowledge provides the guiding principles for effecting change. The second component is a *translational and implementational model*. It converts theoretical principles into an innovative operational model and specifies the content, strategies of change, and their mode of implementation. We often do not profit from our theoretical successes because we lack effective translational and social

diffusion models to disseminate proven psychosocial approaches. The knowledge languishes in our periodicals.

One morning I received a call from Miguel Sabido, a creative dramatist and producer at Televisa in Mexico City. He explained that he is creating engrossing serial dramas, founded on our modeling principles from the Bobo doll studies, to promote literacy and family planning practices nationally. He had devised a remarkably innovative translational model that was producing impressive society-wide changes. We had an informative theory and a creative translational model. But we lacked the expertise and resources to disseminate it widely. The third component, a *social diffusion model* on how to promote adoption of psychosocial programs in diverse cultural milieus, was the missing link. David Poindexter, the director of Population Communications International, heard about the success of this model and became the global diffuser to diverse cultural milieus. Everett Rogers, the foremost theorist in the field of diffusion of innovation, was also an influential player in this global effort. Because social cognitive theory encompasses social network factors in the spread of social influence, this was a synergistic collaboration (Bandura, 1986, 2006b).

These are not social programs foisted on nations by outsiders in pursuit of their self-interest. The dramatic serials are created only on invitation by countries seeking help with intractable problems. Social cognitive [p. 370 ↓] principles are generalizable but their application has to be tailored to the cultural practices and the types of desired changes. This requires functional adaptations. The diffusion center works in partnership with personnel in the host countries to create serial dramas appropriate to their culture and address their particular needs.

In the formative phase, extensive cultural and value analyses are conducted to identify the problems of major concern, the impediments people face, and the improvements they seek in their lives. The dramatizations are grounded in the internationally endorsed values codified in United Nations covenants and resolutions. These values embody respect for human dignity, equitable opportunities, and social practices that support common human aspirations.

Many worldwide applications of this model in Africa, Asia, and Latin America are promoting personal and society-wide changes that are bettering the lives of millions of

people (Bandura, 2006c; Rogers, et al., 1999; Vaughan et al., 2000). They are raising national literacy, enhancing the status of women, reducing unplanned childbearing that contribute to soaring population growth and perpetuate the cycle of poverty, impeding child trafficking for inhumane labor, curtailing the spread of the AIDS epidemic, preserving biodiversity, promoting environmental conservation practices, and in other ways bettering of people's lives.

These global applications illustrate how the power of psychosocial knowledge can be amplified by blending different types of expertise that no one discipline can provide. We often cite examples in the physical and biological sciences where knowledge pursued for its own sake has unforeseen human benefits. The knowledge gained from the Bobo doll experiments conducted years ago spawned unimagined global applications designed to alleviate some of the most pressing global problems. That's a pretty good shelf life.

As I reflect on my journey at this octogenarian milepost, I am reminded of the saying that it is not the miles traveled but the amount of tread remaining that is important. When I last checked, I still have too much tread left to gear down or to bring this engaging psychological Odyssey to a close. In a decision that take goodly control of one's life have begun work on a book o the role of moral disengagement in the exercise of moral agency.

References

- Baer, D.M., Peterson, R.F. and Sherman, J.A. (1967) The development of imitation by reinforcing behavioural similarity to a model . Journal of the Experimental Analysis of Behaviour , 10, 405–416. <http://dx.doi.org/10.1901/jeab.1967.10-405>
- Bandura, A. (1961) Psychotherapy as a learning process . Psychological Bulletin , 58, 143–159. <http://dx.doi.org/10.1037/h0040672>
- Bandura, A. (1965) Vicarious processes: A case of no-trial learning . In L. Berkowitz (ed.) Advances in Experimental Social Psychology , 2, 1–55. New York: Academic Press.

Bandura, A. (1969) *Principles of Behavior Modification* . New York: Holt, Rinehart and Winston.

Bandura, A. (ed.) (1971a) *Psychological Modeling: Conflicting Theories* . New York: Aldine-Atherton.

Bandura, A. (1973) *Aggression: A Social Learning Analysis* . Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1976) Self-reinforcement: Theoretical and methodological considerations . *Behaviorism* , 4, 135–155.

Bandura, A. (1978) The self system in reciprocal determinism . *American Psychologist* , 33, 344–358. <http://dx.doi.org/10.1037/0003-066X.33.4.344>

Bandura, A. (1983) Temporal dynamics and decomposition of reciprocal determinism . *Psychological Review* , 90, 166–170. <http://dx.doi.org/10.1037/0033-295X.90.2.166>

Bandura, A. (1984) Representing personal determinants in causal structures . *Psychological Review* , 91, 508–511. <http://dx.doi.org/10.1037/0033-295X.91.4.508>

Bandura, A. (1986) *Social Foundations of Thought and Action: A Social Cognitive Theory* . Englewood Cliffs: Prentice-Hall.

Bandura, A. (1991a) Social cognitive theory of moral thought and action . In W.M. Kurtines, ed. and J.L. Gewirtz (eds), *Handbook of Moral Behavior and Development*, Vol. 1 . Hillsdale, NJ: Erlbaum.

Bandura, A. (1991b) Human agency: The rhetoric and the reality . *American Psychologist* , 46, 157–162. <http://dx.doi.org/10.1037/0003-066X.46.2.157>

Bandura, A. (1991c) Self-regulation of motivation through anticipatory and self-reactive mechanisms . In R.A. Dienstbier (ed.), *Perspectives on Motivation: Nebraska Symposium on Motivation* , 38, 69–164. Lincoln: University of Nebraska Press.

Bandura, A. (ed.) (1995a) *Self-efficacy in Changing Societies*. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511527692>

Bandura, A. (1995b) Comments on the crusade against the causal efficacy of human thought. *Journal of Behavior Therapy and Experimental Psychiatry*, 26, 179–190. <http://dx.doi.org/10.1016/0005-7916%2895%2900034-W>

Bandura, A. (1996) Ontological and epistemological terrains revisited. *Journal of Behavior Therapy and Experimental Psychiatry*, 27, 323–345. <http://dx.doi.org/10.1016/S0005-7916%2896%2900049-3>

Bandura, A. (1997) *Self-efficacy: The Exercise of Control*. New York: Freeman.

Bandura, A. (1999a) Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, 3, 193–209. http://dx.doi.org/10.1207/s15327957pspr0303_3

Bandura, A. (1999b) A social cognitive theory of personality. In L. Pervin, ed. and O. John (eds), *Handbook of Personality*, 2nd Edition, pp. 154–196. New York: Guilford Publications.

Bandura, A. (2002a) Growing primacy of human agency in adaptation and change in the electronic era. *European Psychologist*, 7, 2–16. <http://dx.doi.org/10.1027//1016-9040.7.1.2>

Bandura, A. (2002b) Social cognitive theory in cultural context. *Journal of Applied Psychology: An International Review*, 51, 269–290. <http://dx.doi.org/10.1111/1464-0597.00092>

Bandura, A. (2004a) Selective exercise of moral agency. In T.A. Thorkildsen, ed. and H.J. Walberg (eds), *Nurturing Morality*, pp. 37–57. New York: Kluwer Academic.

Bandura, A. (2004b) Swimming against the mainstream: The early years from chilly tributary to transformative mainstream. *Behavioral Research and Therapy*, 42, 613–630. <http://dx.doi.org/10.1016/j.brat.2004.02.001>

Bandura, A. (2004c) Health promotion by social cognitive means . *Health Education and Behavior* , 31, 143–164. <http://dx.doi.org/10.1177/1090198104263660>

Bandura, A. (2004d) The role of selective moral disengagement in terrorism and counterterrorism . In F.M. Mogahaddam, ed. and A.J. Marsella (eds), *Understanding Terrorism: Psychological Roots, Consequences and Interventions* , pp. 121–150. Washington, DC: American Psychological Association Press. <http://dx.doi.org/10.1037/10621-006>

BanduraA. (2005a) The evolution of social cognitive theory . In K.G. Smith, ed. and M.A. Hitt (eds) *Great Minds in Management* , pp. 9–35, Oxford: Oxford University Press.

Bandura, A. (2005b) The primacy of self-regulation in health promotion . *Applied Psychology: an International Review* , 54, 245–254. <http://dx.doi.org/10.1111/j.1464-0597.2005.00208.x>

Bandura, A. (2006a) Toward a psychology of human agency . *Perspectives on Psychological Science* , 1, 164–180. <http://dx.doi.org/10.1111/j.1745-6916.2006.00011.x>

Bandura, A. (2006b) On integrating social cognitive and social diffusion theories . In A Singhal, ed. and J. Dearing (eds), *Communication of Innovations: A Journey With Ev Rogers* , pp. 111–135. Beverley Hills: Sage Publications.

Bandura, A. (2006c) Going global with social cognitive theory: From prospect to paydirt . In S.I. Donaldson, ed. , D.E. Berger, ed. , and K. Pezdek (eds), *Applied Psychology: New Frontiers and Rewarding Careers* , pp. 53–79. Mahwah, NJ: Lawrence Erlbaum.

Bandura, A. (2006d) Guide to the construction of self-efficacy scales . In F. Pajares, ed. and T. Urdan (eds), *Self-efficacy Beliefs of Adolescents* , 5, 307–337. Greenwich, CT: Information Age Publishing.

Bandura, A. (2008a) The reconstrual of ‘free will’ from the agentic perspective of social cognitive theory . In J. Baer, ed. , J.C. Kaufman, ed. and R.F. Baumeister (eds), *Are We Free? Psychology and Free Will* , pp. 86–127. Oxford: Oxford University Press.

Bandura, A. (2008b) An agentic perspective on positive psychology . In S.J Lopez (ed.), Positive Psychology: Exploring the Best in People , 1, 167–196. Westport, CT: Greenwood Publishing Company.

Bandura, A. (2009) Social cognitive theory of mass communications . In J. Bryant, ed. and M.B. Oliver (eds), Media Effects: Advances in Theory and Research , 2nd Edition, pp. 94–124. Mahwah, NJ: Lawrence Erlbaum.

Bandura, A., Blanchard, E.B. and Ritter, B. (1969) Relative efficacy of desensitization and modeling approaches for inducing behavioral, affective, and attitudinal changes . Journal of Personality and Social Psychology , 13, 173–199. <http://dx.doi.org/10.1037/h0028276>

Bandura, A. and Locke, E. (2003) Negative self-efficacy and goal effects revisited . Journal of Applied Psychology , 88, 87–99. <http://dx.doi.org/10.1037/0021-9010.88.1.87>

Bandura, A., Ross, D. and Ross, S.A. (1963) Imitation of film-mediated aggressive models . Journal of Abnormal and Social Psychology , 66, 3–11. <http://dx.doi.org/10.1037/h0048687>

Bandura, A., Underwood, B. and Fromson, M.E. (1975) Disinhibition of aggression through diffusion of responsibility and dehumanization of victims . Journal of Research in Personality , 9, 253–269. <http://dx.doi.org/10.1016/0092-6566%2875%2990001-X>

Bandura, A. and Walters, R.H. (1963) Social Learning and Personality Development . New York: Holt, Rinehart and Winston.

Boffey, P. and Walsh, J. (1970) Study of TV violence: Seven top researchers blackballed from panel . Science , 168, 949–952. <http://dx.doi.org/10.1126/science.168.3934.949>

Braithwaite, J. (1994) A sociology of modeling and the politics of empowerment . British Journal of Sociology , 45, 445–479. <http://dx.doi.org/10.2307/591658>

- Carver, C.S. and Scheier, M.F. (1981) Attention and Self-regulation: A Control-theory Approach to Human Behavior . New York: Springer-Verlag. <http://dx.doi.org/10.1007/978-1-4612-5887-2>
- Catania, C.A. (1975) The myth of self-reinforcement . *Behaviorism* , 3, 192–199.
- Cater, D. and Strickland, S. (1975) TV Violence and the Child . New York: Russell Sage Foundation.
- Ephron, E. (1969a) A Child Is Not A Rat . *TV Guide* , November 26.
- Efron, E. (1969b) The man in the eye of the hurricane . *TV Guide* , November 15.
- Film Board of Canada (1972) The Question of Violence: The US Senate Hearings . [Motion Picture]. United States: Phoenix Films. (Available from Phoenix Films, Inc., 470 Park Avenue South, New York, NY 10016).
- Gewirtz, J.L. (1971) Conditional responding as a paradigm for observational, imitative learning and vicarious-reinforcement . In H.W. Reese (ed.), *Advances in Child Development and Behavior* , 6, 273–304. New York: Academic Press.
- Goldsen, R. (1972) Science in wonderland . Unpublished manuscript, Cornell University, Ithaca, NY.
- Gould, J. (1972) Violence Held Unharmful to Youth . *New York Times* , p. 27.
- Gould, S.J. (1987) An Urchin in the Storm . New York: Norton.
- Haney, C. (1997) Violence and the capital jury: Mechanisms of moral disengagement and the impulse to condemn to death . *Stanford Law Review* , 49, 1447–1486. <http://dx.doi.org/10.2307/1229350>
- Harré, R. (1983) Personal Being: A Theory for Individual Psychology . Oxford: Blackwell.
- James, W. (1884) Absolutism and empiricism . *Mind* , 9, 281–286. <http://dx.doi.org/10.1093/mind/os-IX.34.281>

Kim, U., Triandis, H.D., Kâgitçibasi, C., Choi, S. and Yoon, G. (1994) Individualism and Collectivism: Theory, Method, and Applications . Thousand Oaks, CA: Sage.

Locke, E.A. (1994) The emperor is naked . International Review of Applied Psychology , 43, 367–370. <http://dx.doi.org/10.1111/j.1464-0597.1994.tb00829.x>

Lord, R.G. and Levy, P.E. (1994) Moving from cognition to action: A control theory perspective . Applied Psychology: An International Review , 43, 335–398. <http://dx.doi.org/10.1111/j.1464-0597.1994.tb00828.x>

McCrae, R.R. and Costa, P.T. (1996) Toward a new generation of personality theories: Theoretical contexts for the five-factor model . In J.S. Wiggins (ed.), Five-factor model of personality: Theoretical perspectives , 51–87. New York: Guilford Press.

Milgram, S. and Shotland, R.L. (1973) Television and Antisocial Behavior: Field Experiments . New York: Academic Press.

Miller, N.E. and Dollard, J. (1941) Social Learning and Imitation . New Haven: Yale University Press.

Muñoz, R. and Mendelson, T. (2005) Toward evidence-based interventions for diverse populations: The San Francisco General Hospital prevention and treatment manuals . Journal of Consulting and Clinical Psychology , 73, 790–799.

Nagel, E. (1961) The Structure of Science . New York: Harcourt, Brace and World.

National Commission on the Causes and Prevention of Violence (1969) Commission statement on violence in television entertainment programs . Washington, DC: US Government Printing Office.

Paisley, M. (1972) Social Policy Research and the Realities of the System: Violence Done to TV Research . Stanford, CA: Stanford University, Institute for Communication Research.

Patterson, G.R. (1982) A Social Learning Approach (Vol 3). Coercive family process . Eugene, OR: Castalia.

Phillips, D.C. and Orton, R. (1983) The new causal principle of cognitive learning theory: Perspectives on Bandura's 'reciprocal determinism' . *Psychological Review* , 90, 158–165. <http://dx.doi.org/10.1037/0033-295X.90.2.158>

Powers, W.T. (1973) *Behavior: The Control of Perception* . Chicago: Aldine.

Powers, W.T. (1978) Quantitative analysis of purposive systems: Some spadework at the foundations of scientific psychology . *Psychological Review Monograph Supplements* , 85, 417–435.

Powers, W. T. (1991) Comment on Bandura's 'human agency' . *American Psychologist* , 46, 151–153. <http://dx.doi.org/10.1037/0003-066X.46.2.151.b>

Rogers, E.M., Vaughan, P.W., Swalehe, R.M.A., Rao, N., Svenkerud, P. and Sood, S. (1999) Effects of an entertainment-education radio soap opera on family planning behavior in Tanzania . *Studies in Family Planning* , 30, 1193–1211. <http://dx.doi.org/10.1111/j.1728-4465.1999.00193.x>

Rosenthal, T.L. and Zimmerman, B.J. (1978) *Social Learning and Cognition* . New York: Academic Press.

Rotter, J.B. (1954) *Social Learning And Clinical Psychology* . Englewood Cliff, NJ: Prentice- Hall. <http://dx.doi.org/10.1037/10788-000>

Sampson, E.E. (1977) Psychology and the American ideal . *Journal of Personality and Social Psychology* , 35, 767–782. <http://dx.doi.org/10.1037/0022-3514.35.11.767>

SchoolerC. (1990) Individualism and the historical and social-structural determinants of people's concerns over self-directedness and efficacy . In J. Rodin, ed. , C. Schooler, ed. and K.W. Schaie (eds), *Self-directedness: Cause and Effects Throughout the Life Course* , pp. 19–58. Hillsdale, NJ: Erlbaum.

Skinner, B. F. (1971) *Beyond Freedom and Dignity* . New York: Knopf.

Skinner, B.F. (1977) Why I am not a cognitive psychologist . *Behaviorism* , 5, 1–10.

Staddon, J.E.R. (1984) Social learning theory and the dynamics of interaction . Psychological Review , 91, 502–507. <http://dx.doi.org/10.1037/0033-295X.91.4.502>

Stoltz, S.B. (1978) Ethical Issues in Behavior Modification . San Francisco: Jossey-Bass.

Vaughan, P.W., Roger, E.M., Singhal, A. and Swalehe, R.M. (2000) Entertainment-education and HIV/AIDS prevention: A field experiment in Tanzania . Journal of Health Communications , 5, 81–100. <http://dx.doi.org/10.1080/10810730050019573>

Wilson, E.O. (1998) Consilience: The Unity of Knowledge . New York: Knopf.

Author's Note

Some sections of this chapter include revised, updated, and elaborated material from the following publications:

Bandura, A. (2005) *The evolution of social cognitive theory*. In K.G. Smith, ed. and M.A. Hitt (eds) *Great Minds in Management*, pp. 9–35. Oxford: Oxford University Press.

Bandura, A. (2006) *Toward a psychology of human agency. Perspectives on Psychological Science*, 1, 164–180.

Bandura, A. (2006) *Autobiography*. M.G. Lindzey, ed. and W.M. Runyan (eds) *A History of Psychology in Autobiography IX*, 42–75. Washington, DC: American Psychological Association.

Chapters and research articles that address issues discussed in this publication are available in electronic form in the following website: <http://des.emory.edu/mfp/Bandura>.

[p. 374 ↓]

<http://dx.doi.org/10.4135/9781446249215.n18>

Handbook of Theories of Social Psychology: Volume 1

Cognitive Dissonance Theory

Contributors: Joel Cooper

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Cognitive Dissonance Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n19>

Print pages: 377-398

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n19>

[p. 377 ↓]

Chapter 18: Cognitive Dissonance Theory

Joel Cooper, ed.

Abstract

Cognitive dissonance theory has been a major pillar of social psychology for decades. In this chapter, I discuss some of the reasons that Festinger's straightforward proposition about the relationship among cognitions created the pointed controversy that propelled it into the forefront of the discipline. Theories with the broad approach of cognitive dissonance often need modification and this chapter traces that journey. I then present the New Look and Self-Standards models that attempted to integrate extant data and changed our understanding of the motivational foundation that underlies dissonance. Current perspectives on cognitive dissonance focused on vicarious dissonance arousal by which individuals experience dissonance based on the behavior of fellow group members. Finally, the chapter examines the potential use of personal cognitive dissonance for optimizing the effectiveness of psychotherapy and the use of vicarious dissonance for increasing positive behaviors to protect health and wellbeing.

In the Beginning

The theory of cognitive dissonance has been a staple of social psychology for more than half a century. In this chapter, I will present my own perspective on the birth of dissonance and then take the responsibility for two alternative approaches that modified our view of the meaning of cognitive dissonance: the New Look model (Cooper and Fazio, 1984) and the subsequent Self-Standards model of dissonance (Stone

and Cooper, 2001). The story begins with Festinger's observations about cognitive inconsistency and will move to our newer views of the motivation underlying cognitive dissonance.

The late Leon Festinger molded the original theory of cognitive dissonance from his interest in people's susceptibility to field forces (Lewin, 1951), including pressures from groups. He had published a number of major statements about the pressures that groups place on individuals to achieve attitudinal consensus (Festinger, 1950). In 1954, he shifted his focus to the study of the individual. Instead of viewing pressure from the vantage point of the group's needs and goals, he took the perspective of the individual who was driven to use others as a benchmark to measure his or her own standing in a group. He proposed that people were driven to [p. 378 ↓] compare their opinions and abilities with similar others and that there was pressure either to conform to the attitudes of similar others or to convince others to hold attitudes similar to oneself (Festinger, 1954).

In the theory of cognitive dissonance, Festinger (1957) completed the task of viewing the world from the perspective of the individual. In dissonance theory, cognitive consistency was represented inside the head of the person. To think of mental life as a set of cognitive representations was a radical departure from the mainstream view of the 1950s. For the first time, people's views of their social world, their appraisals of their fellow group members, their own opinions about the world, and their observations of their own and others' behaviors could all be projected onto a common screen. All were cognitive representations inside the head. Moreover, some of those cognitive representations bore a relationship to each other. The birth of cognitive dissonance theory occurred at that instant. The new theory – cognitive dissonance – became the most productive of all of his creative insights, and dissonance theory charted a research agenda that would last for half a century.

Cognitive Dissonance in a Learning Theory World

The major tenets of the original version of dissonance theory are well known and straightforward. The state of cognitive dissonance occurs when people perceive that a pair of cognitions is inconsistent. Formally, Festinger defined a pair of cognitions as dissonant if the actor believed that one cognition followed from the obverse of the other. He postulated that dissonance is experienced as an unpleasant drive and, like other unpleasant drive states, needs to be reduced. The reduction occurs by changing the cognition least resistant to change or by adding cognitions that minimize the perceived magnitude of the discrepancy. In keeping with Festinger's philosophical assumption that the dissonance battle was played out inside the head of the perceiver, he reasoned that inconsistency itself is a psychological state – that is, two cognitions are dissonant if the perceiver *believes* they are dissonant. The psychology of the perceiver, not the philosophical rules of logic, determines the existence of dissonance.

The idea that people prefer consistency to inconsistency was not new. Fritz Heider (1946) and Theodore Newcomb (1956) had written about such ideas previously and they were also consistent with field theoretical notions of Festinger's advisor, Kurt Lewin. To my knowledge, Festinger's 1957 book outlining his ideas did not raise controversy until the publication two years later of Festinger and Carlsmith's (1959) now classic study showing that people experienced dissonance after an attitude-inconsistent statement. As we know, Festinger and Carlsmith had people participate in a task that was specifically designed to be tedious and dull. Those participants then agreed to make a statement to a person, who they believed was a fellow student, extolling the excitement of the task. Few would have had difficulty with the finding that making the statement about the excitement of the task induced people to change their attitudes in the direction of their statement. If we accept the premise that people do not like inconsistency and that they are motivated to reduce the incongruence between their behavior and attitude, then the finding makes sense and would have been predicted by any consistency theory.

The provocative element in Festinger and Carlsmith's (1959) research was the role played by the magnitude of the incentive that participants were offered to make a public statement that was contrary to their beliefs. Some were offered a large incentive while others were offered a meager incentive. Festinger and Carlsmith reasoned that the large incentive would be a sufficiently important cognition consistent with the behavior to keep dissonance low, whereas people offered a small incentive would still be in the throes [p. 379 ↓] of dissonance. The finding that behavior associated with small incentives could create more change than behavior associated with large incentives had the effect of prodding a large sleeping animal with a small stick. It woke the animal and, to continue the metaphor, allowed the provocateur not only to be noticed but, in the end, to assume a leadership position in the metaphorical jungle.

The year 1957 was the realm of learning theories. The number of people considering themselves social psychologists was scant. The "science" of psychology was focused on the rules of sensation, perception, and learning. The last of these topics was particularly vibrant, with lively debates among followers of Hull, Spence, Tolman, and Skinner filling the pages of the professional literature. They disagreed about many issues such as the importance of habits and the proper role of drive states. What they all agreed upon, however, was the role of rewards and reinforcements. Although they conceptualized them differently, larger rewards led to more behavior change; smaller rewards led to less change. This was the gospel according to learning theory.

Dissonance theory, and Festinger and Carlsmith's findings in particular, put the brakes on that assumption. The playing field changed. Suddenly, attitudinal and behavioral change were at the service of smaller rewards rather than larger ones. Large incentives merely reduced the dissonance drive state and led to less change than smaller rewards, or perhaps no rewards at all. We should also not underestimate Festinger's use of the drive state as the motivation for change. By positing that people were motivated by what was essentially a drive, he positioned dissonance theory alongside the major learning theories in which drive-reduction played the critical role. It is unclear whether Festinger ever believed that we would ever find evidence for the drive concept, but by using it as the motivational metaphor, his findings were instantly recognized as a challenge to all who wanted to see social behavior as merely a carryover of the behavioral rules that applied to rats and pigeons.

Other findings continued to goad the conventional learning theory wisdom. The more people suffered, the more they liked what they suffered to attain (Aronson and Mills, 1959). The less punishment with which children were threatened to refrain from playing with an attractive toy, the more they devalued the toy (Aronson and Carlsmith, 1963) and the longer they refrained from playing with it (Freedman, 1965). Learning theorists rose to the challenge of these findings, criticizing the methodology, the conclusions, and the theory. Journals, including the new *Journal of Personality and Social Psychology*, were filled with debates about the intriguing new ideas predicted from cognitive dissonance theory. The early controversial years of provocative dissonance results were followed by challenges from critics, and followed again by studies from a growing number of dissonance researchers. By the early 1970s, dissonance had become a “movement” and very few questioned the existence of dissonance as a powerful principle of human social behavior.

Trouble in Dissonance Heaven

My own activity as a dissonance researcher came from an unlikely source. I attended graduate school at Duke University, partly to study with Jack W. Brehm, one of Festinger's original students and the researcher responsible for the very first published experiment based on dissonance theory (Brehm, 1956). However, Jack was busy with his new theory of psychological reactance and I was assigned to work with Edward E. Jones, who was soon to become a towering figure in the world of person perception and attribution. Nonetheless, Jones was a fan of dissonance research and was perplexed by a study that had been published by Rosenberg (1965) that seemed to spell trouble for the fledgling theory.

[p. 380 ↓] Rosenberg had argued that the reason for Festinger and Carlsmith's startling result was that participants had been worried that they were being evaluated on their level of psychological consistency. He coined the term *evaluation apprehension* to capture participants' concern that the person who had asked them to say that the task was fun was also privy to their attitudinal confessions when they responded to the dependent measure. Rosenberg provided evidence that attitude change was a direct function of level of reward (i.e., more attitude change for higher compensation) when the experimenter who collected the dependent measure was not in any way connected to

the person who made the request for the counterattitudinal behavior. Jones suspected that the key element missing in Rosenberg's critical study was the element of choice. Although Festinger had not paid much attention to the need for a person to *choose* to engage in counterattitudinal behavior, he had built it into his original experiment (Festinger and Carlsmith, 1959). So, my advisor sent me out to devise a study to show that how people perceived their freedom to engage in attitude discrepant behavior made all of the difference in the arousal of dissonance.

The results of our study were published as Linder, Cooper, and Jones (1967) – my very first published dissonance article. In two experiments, we showed that attitude change can be a direct function of incentive magnitude or an inverse function, depending on whether dissonance is ever aroused by the procedure. If people perceive that they were free to accept or decline the invitation to make a counterattitudinal statement, then dissonance is aroused. We found that when people had been given the freedom to decline, then just like Festinger and Carlsmith had found, attitudes were an inverse function of incentive magnitude: the lower the reward, the greater the attitude change. It is only in the absence of dissonance – brought about by the absence of decision freedom – that attitude change became a direct function of the magnitude of the reward. In other words, reinforcement works to elicit attitude change but is trumped by dissonance arousal. When dissonance exists – in this case, facilitated by the perception of decision freedom – it provides the underpinning for attitude change following counterattitudinal advocacy. (Later, Jones told me that I got his idea completely wrong, ran a different experiment from the one he thought I was running, but had no complaints about the outcome!)

The Theory Evolves: The Search for Modifiers Becomes the New Look Model of Dissonance

A broad-stroke theory almost always needs modifiers. Many of the studies in the first decade of dissonance theory did exactly that. Linder et al.'s (1967) results were of that genre. Dissonance is aroused and leads to attitude change when decision freedom is

high but not when it is low. We thus learned that choice was a modifier of dissonance. Carlsmith et al. (1966) as well as Davis and Jones (1960) had shown that people need to feel committed to their counterattitudinal behavior. If participants thought they would have a chance to “take back” their counterattitudinal statements, then dissonance would not occur. We thus learned that commitment was a modifier of cognitive dissonance.

A few years later, Steve Worchel and I wondered aloud whether *any* counterattitudinal statement expressed in *any* circumstance produced dissonance or whether something had to happen as a result of the behavior. Put another way, would an attitude discrepant statement uttered in the dark of one's own residence with no one to hear the remark cause the arousal of dissonance? We suspected that the solitary utterance would not lead to dissonance. We designed a study that replicated the procedure of Festinger and Carlsmith's (1959) with the added factor of whether or not the student in the waiting room seemed convinced by the participant's statement that the task was fun and interesting. We predicted and found that the inverse [p. 381 ↓] relationship between incentive magnitude and attitude change following counterattitudinal behavior occurred only when there was a consequence pursuant to the behavior – in this case, misleading a fellow student to have a false expectation (Cooper and Worchel, 1970). We had found another modifier: inconsistent cognitions lead to dissonance in the presence, but not the absence, of an unwanted consequence.

In 1980, a visiting professor at Princeton University, Paul Secord, expressed his consternation about the state of play of dissonance theory. He told me that he loved the theory when it was broad, simple, and comprehensible – that is, inconsistent cognitions lead to the state of dissonance. However, with two decades of research under the dissonance tent, he felt he needed a scorecard to know when it is true that inconsistent cognitions arouse dissonance and when it is not. He expressed the thought that the modifiers of dissonance had become the face of the theory and that someone needed to make sense of the modifiers. That was the challenge that my former graduate student, Russell Fazio, and I took up in the paper that became our signature modification of dissonance theory: the “new look” (Cooper and Fazio, 1984).

The New Look Model of Dissonance

Fazio and I examined the published data. Secord had a point. Inconsistent cognitions aroused dissonance

- *but only if the actions were freely chosen;*
- *but only if the actor was committed to the discrepant cognition;*
- *but only if there was an aversive event following the dissonance;*
- *but only if the aversive consequences were foreseeable.*

And the list went on. The theory was sorely in need of a cure for the *but onlys*.

It occurred to us that the important quest to find the modifiers of dissonance had contributed to a more meaningful understanding of dissonance. Looked at from a slightly different perspective, the research of the prior two decades had transformed dissonance into a different kind of theory. Rather than maintaining that dissonance is a function of inconsistency, the data had actually called out for a new statement about the meaning of dissonance. The data were there. The *New Look* just had to tell the new story in different words.

As Festinger had surmised, dissonance is an arousing, uncomfortable tension state that motivates change. However, it is not brought about by cognitive inconsistency per se, but rather by the perception that one is responsible for bringing about an unwanted event. That perception, rather than the perception of inconsistency, is what results in the experience of cognitive dissonance.

Dissonance Arousal and Dissonance Motivation

According to the New Look, dissonance begins with a behavior. In order for that behavior to lead to a cognitive or attitude change, a set of processes must unfold that can best be divided into two stages: *dissonance arousal* and *dissonance motivation*.

Dissonance arousal occurs when people take responsibility for bringing about an aversive event. This conclusion may come about quickly, but not easily. There are several decision points that need to be crossed in order for a behavior to bring about dissonance arousal.

First, a behavior has to be perceived to have an unwanted consequence. Almost every behavior has a consequence. The question at issue for the actor is whether the consequence is unwanted and, if so, how unwanted is it? For example a person may favor a particular solution to the healthcare dilemma in the United States but advocate for a bill that is a bit different. If the person successfully convinces her friends, colleagues, or senators to support the bill, is this consequence sufficiently unwanted to lead to dissonance? Here, we argued that the consequence has to [p. 382 ↓] fall outside a person's "latitude of acceptance" of possible positions that she can accept in order for dissonance to be aroused, otherwise dissonance will not follow from the behavior. Empirical evidence supports this first link in the dissonance process (Fazio et al. 1977).

The second decision point is the acceptance of *personal responsibility* for the consequences of the behavior. We defined responsibility as a combination of two factors: freely choosing the behavior in question and the ability to foresee the consequences of that behavior. Accepting responsibility leads to dissonance; denial of responsibility allows people to avoid the unpleasant state of dissonance. Decision freedom is crucial because it is necessary for the acceptance of responsibility.

Although decision freedom is necessary, it is not sufficient to lead to acceptance of responsibility. Imagine that the healthcare advocate in our example purchased a book that examined the potential uses and abuses of privatized healthcare. After the purchase, she found out that the proceeds from the book were donated to an advocacy group that condemned any government-supported healthcare plan. Would our advocate feel personally responsible for supporting a group that she finds unacceptable and contrary to her values? The New Look model argues that the answer is no. Despite the fact that she really gave money in support of the disliked group, she had no way to foresee that her actions would lead to the unwanted consequence. Foreseeability, then, is the second element that combines with decision freedom to determine whether one is responsible for an aversive event (Goethals et al., 1979). It is the acceptance

of personal responsibility that provides the necessary and sufficient link to dissonance arousal.

On the Radical Nature of the New Look

We did not intend the New Look to be a radical departure from Festinger's original notion, but we soon saw that it was. The arousal of dissonance no longer depended on what had been its major tenet – that is, the presence of inconsistency. It is true that inconsistency is typically a reasonable proxy for the active ingredients that lead to dissonance arousal. When people behave in ways that are inconsistent with cherished beliefs, they are typically simultaneously taking personal responsibility for bringing about at least the potential for an unwanted event. They may never find out if someone will be convinced by an attitude-discrepant statement, but the potential for bringing about an unwanted event is apparent. Thus, according to the New Look, acting inconsistently often brings with it the features that are actually the active ingredients in the dissonance process. Similarly, when people choose one attractive item over another in the free-choice research paradigm (Brehm, 1956) of dissonance, they are responsible for the aversive consequence: that is, rejecting all of the attractive elements of the unchosen alternative and accepting all of the unattractive features of the chosen alternative. And when people in an effort justification study (Aronson and Mills, 1959) choose to suffer in order to attain a goal, they are responsible for engaging in the unpleasant work, embarrassment or effort that is, in itself, an unwanted consequence. The problem with relying on inconsistent cognition as the underpinning of cognitive dissonance is that it requires a long list of caveats and exceptions (the *but onlys*) that moderate the effect. The underpinnings of the New Look model permit a more comprehensive view of the process.

Testing the New Look

Most of the work on the New Look model of dissonance preceded the publication of the New Look. As explained, the New Look was a way of understanding the theoretical underpinnings of dissonance, given what we already knew from published work

on dissonance. [p. 383 ↓] I would postulate that there were four major categories of findings on which the New Look relied. First, research must show that people experience dissonance arousal when they choose to engage in attitude-discrepant behaviors and not when they were forced to do so. This finding had already been well established prior to the New Look and continues to be a reliable research result. Second, attitude discrepant behavior must be shown to lead to dissonance when and only when the potential for aversive consequences exists. We had already shown this phenomenon on many occasions prior to the New Look model. Cooper and Worchel (1970) were the first to demonstrate it, but replications of the effect abound (Cooper et al., 1974; Goethals and Cooper, 1972; Goethals et al., 1979; Norton et al., 2003). (In fairness, this feature has emerged as a more controversial issue with some research questioning not its importance but its ubiquity [Harmon-Jones et al., 1996]). A third critical feature of the model was that dissonance arousal relies on the consequences of a behavior being foreseeable at the time they commit to their behavior. This, too, was well established (Cooper and Goethals, 1974; Goethals and Cooper, 1972; Goethals et al. 1979).

The fourth crucial feature had not been demonstrated when the model was first introduced. The responsibility-for-consequences model predicts that any behavior – not just inconsistent behavior – will arouse cognitive dissonance if it foreseeably leads to an aversive event. The radical nature of the New Look lay in its position that inconsistency was not necessary for dissonance to occur. In a study designed to collect evidence for this hypothesis, Scher and Cooper (1989) had people commit to attitude-consistent or attitude-inconsistent behavior. A cover story led participants to believe that a university committee was considering a policy in which student health records would become open to parents to peruse. Some were asked to write counterattitudinal essays that supported this unpopular and unwanted policy while others were asked to write proattitudinal essays. The students were led to believe that the essays might either convince the committee or might boomerang and convince the committee of the opposite of what was written. In this way, the behavior (counterattitudinal versus proattitudinal) was manipulated orthogonally with the side of the issue that the committee was likely to believe (a desired versus an unwanted consequence of the essay).

When attitudes were measured following the essay, the results showed an effect for the consequence of the essay and not for the discrepancy of the arguments. Counterattitudinal advocacy led to attitude change only if it had the potential to produce an aversive event, thus replicating the finding of Cooper and Worchel (1970). But so did proattitudinal advocacy. Writing in the direction of their own beliefs, students who discovered that their essays were likely to produce a boomerang and thus help to bring about an unwanted event changed their attitudes. The consequence and not the inconsistency of behavior and attitude produced dissonance arousal that ultimately motivated attitude change.

Is Dissonance Arousing?

One of the hallmarks of Festinger's brilliance was his adoption of the drive model to introduce the theory of cognitive dissonance. Clark Hull, Kenneth Spence and their colleagues were debating the proper mix of drives and habits to understand how organisms learn. M.L. Skinner (1953) had made the unique proposal that drives were not necessary at all to understand learning. The drive concept was highly accessible and controversial as Festinger introduced cognitive dissonance to the community.

By postulating a drive state underlying the dissonance process, he made his work relevant to the learning community both inside and outside of social psychology. By postulating that dissonance reduction in social behavior is often an inverse function of the [p. 384 ↓] magnitude of incentive and reward, he upset the foundation that the learning community had laid for social psychologists. I believe that the drive and reinforcement notions were both key to the influence that dissonance had in the 1950s. I was never certain that Festinger ever expected the drive concept to receive a direct test. It was a virtual metaphor, a way to think about the process that had the added benefit of making his fledgling theory both noticed and controversial. I believe he was pleasantly surprised when subsequent research found that he had guessed correctly.

Converging evidence has now demonstrated that dissonance has, as Festinger (1957) called it, "drive-like properties." Waterman and Katkin (1967) reasoned that if dissonance was a drive, then it should have the effect that drives typically have on learning: it should facilitate simple learning and interfere with complex learning.

Waterman and Katkin (1967) found evidence for the former but not for the latter. That evidence was supplied a few years later by Pallak and Pittman (1972) who found that dissonance following counterattitudinal advocacy interfered with people's ability to learn a complex task.

Using a different logic, Zanna and Cooper (1974) showed that if people believed their arousal following counterattitudinal advocacy was due to something other than the advocacy (such as a pill), they did not show attitude change following the advocacy. Apparently, attitude change is directed toward lowering the uncomfortable drive state. If the arousal was thought to be due to some other agent, no attitude change ensued. Moreover, reducing the degree of bodily arousal by means of a sedative was shown to decrease attitude change following counterattitudinal advocacy while ingesting an arousing agent was shown to increase attitude change (Cooper et al. 1978)

A third line of evidence was the measurement of arousal following counterattitudinal advocacy. Croyle and Cooper (1983) found skin conductance differences between participants in high dissonance and low dissonance conditions. Losch and Cacioppo (1990) replicated that finding and also showed that dissonance reduction is directed at reducing the uncomfortable affect that dissonance produces. Elliot and Devine (1994) added to the literature on the uncomfortable motivational state by asking participants how they felt following a dissonance-arousing act. Participants reported significantly more negative discomfort than participants in low dissonance conditions.

The New Look complements the original version of dissonance theory by accepting the notion that cognitive changes are motivated by psychological discomfort and physiological arousal. What may have begun as a metaphor to predict change has received substantial support from a variety of research perspectives. Dissonance, we believe, is motivated by being responsible for bringing about an unwanted consequence of behavior. It is experienced as discomfort and motivates cognitive change.

Self-Standards: Moving the New Look Forward

The New Look model of dissonance was not without its critics (see Harmon-Jones and Mills, 1999). Elliot Aronson, one of dissonance theory's most innovative pioneers, advocated a perspective on dissonance that focused on an individual's violating his or her concept of self. Good people expect that they will do good things; bad people expect they will do bad things and when the twain mixes, dissonance is aroused. Early work (e.g., Aronson and Carlsmith, 1962) had shown that people who expect to fail will choose to fail as a way of staying consistent with their self-concept. To do otherwise would result in the unpleasant inconsistency known as dissonance. Aronson (e.g., 1992; Thibodeau and Aronson, 1992) argued that inconsistency with self was sufficient to produce dissonance and the New Look's [p. 385 ↓] insistence on aversive consequences was unnecessary.

In order to demonstrate that dissonance can occur in the absence of aversive events, Thibodeau and Aronson (1992) introduced what they called the 'hypocrisy paradigm'. In this research, people were asked to make statements that were consistent with their private beliefs but were reminded of times in the past that they had acted inconsistently. The general finding from this research is that dissonance occurs as a function of the proattitudinal statement when the statements are made freely and participants are made mindful of their prior inconsistent behavior. Typically, participants engage in subsequent behavior that is more in line with their proattitudinal advocacy. For example, college students in a study by Stone et al. (1994) were asked to make public statements to a group of high school students advocating the use of condoms when engaging in sex. This speech was consistent with the participants' attitudes about the use of condoms. In a key hypocrisy condition, the participants were made mindful of occasions in which they had not practiced what they had just preached – that is, they were asked to recall times that they had not used condoms. When the study was allegedly finished, the participants were allowed to purchase as many condoms as they wished. The results showed that when making the statement under conditions of free choice and with the reminder of their own discrepant behavior, participants purchased more condoms.

than in any other condition in the experiment. In order to reduce their dissonance, participants exaggerated their behavior to bring it in line with their attitude and their proattitudinal statement.

This intriguing line of research raised the question of the necessity of an aversive event for the dissonance process (Aronson, 1992). On the other hand, it seemed to me (Cooper, 1992) that the aversive consequence was intrinsically enmeshed in the mindfulness manipulation. Being reminded of not having worn condoms or, in other studies, being reminded of times that you wasted water (Dickerson et al., 1992) or failed to recycle (Fried and Aronson, 1995) are all aversive consequences. They happen to be in the past, but they are consequences nonetheless. The results seemed consistent with the New Look model.

The set of hypocrisy studies accomplished a great deal, however. It added to the number of ways that researchers can investigate dissonance arousal; the dissonance resulting from hypocrisy is frequently channeled to behavior change rather than the more typical attitude change and the behavior change is usually in a direction that promotes constructive social and personal values – an issue that we will return to later in this chapter. From a theoretical standpoint, the studies laid bare an issue about which we had said little in the New Look model: What do we mean by an aversive consequence?

In the New Look, Fazio and I defined an aversive consequence as an occurrence that one would rather not have occur. That is, if you can think of something you would not like to bring about, such as convincing a person to believe in an unwanted position or suffering embarrassment or being stuck with unattractive features of a choice alternative, then that is what we meant by “aversive consequence.” We believed that there was no adequate *a priori* definition of an aversive consequence. Whatever a person thought was unwanted and yet acted in a way that caused that event to occur was grist for the dissonance mill. We had no disagreement with Aronson's (1969) notion that violations of self-expectancies can cause dissonance – as long as people find it aversive to violate their self-expectancies. The New Look differed from self-expectancy view because we did not believe that self-expectancy violations were the *only* route to dissonance. Whatever a person finds aversive or unwanted, whether it is a violation of self-expectancy or a behavior that brings about any other unwanted consequence

fits the New Look's understanding of an aversive event and serve to arouse cognitive dissonance.

[p. 386 ↓] Jeff Stone, who had been a graduate student with Elliot Aronson and then a postdoctoral fellow working with me, helped to integrate the New Look's emphasis on any aversive consequence and self-expectancy's emphasis on violations of the self as necessary for dissonance. In what I believe to be the most recent full model of dissonance processes, Stone and I (Stone and Cooper, 2001) advanced the Self-Standards model of dissonance. What was missing in the New Look was an explicit way to judge the meaning of a behavior. In Stone and Cooper (2001) we argued that what arouses dissonance is an initial assessment of a behavior against a particular behavioral standard. All behaviors have consequences. To judge the desirability of those consequences requires a comparison to a standard of judgment. In the Self-Standards model we spelled out those judgment standards.

Normative and Personal Standards

We reasoned that there are two major categories of standards that a person can use to assess the meaning of the consequences of his or her behavior – normative and personal. There are some outcomes that we can create in the world that most people would agree are of a particular valence. Most people would agree that contributing to a charity or helping a roommate study for an exam are positive events. We know there may be occasions in which helping a roommate and/or contributing to charity may have complicated mixed motives, but by and large, such actions are considered positive. Similarly, there are consequences that most people would agree are negative or undesired. Running into someone on the street and knocking him down would be generally aversive. So, too, is lying to someone, especially when the person believes you and is influenced by your lie (e.g., Festinger and Carlsmith, 1959). Granted, there may be some odd times when those outcomes are positive, but, typically, most people would agree they are negative.

When a standard of judgment is based on a perception of what *most people* perceive to be foolish, immoral, or otherwise negative, we can say people are using a *normative standard of judgment*. The main thrust of this definition is that the standards are based

on a shared understanding of good and bad, wanted or unwanted, foolish or clever (Higgins, 1989). The other broad category of standards of judgment are those that are based on the unique characteristic of the individual. These are *personal standards of judgment*. They refer solely to the judgments people make when they consider only their own values or desires. Consider a casual runner who runs a mile in 4.5 minutes. By the standards of most casual runners, this is an extraordinary experience. However, this particular runner expected to cross the mile marker in closer to 4 minutes. Regardless of whether it is rational or not, regardless of whether others would agree with this runner's judgment, the outcome fails to meet the runner's personal standard of judgment. The achievement, when compared with the runner's personal standard of judgment, is not an achievement at all but an unwanted, aversive event.

The self-standards dissonance model asserts that people can use either a normative standard of judgment or a personal standard of judgment in order to assess their behavior. Which standard they use is a function of the standard that is accessible at the time of their behavior. If the situation makes normative standards accessible, then people will use their concept of what most people would find desirable as the way to assess the consequences of their behavior. Conversely, if people are induced to have their personal standards accessible, then they will use their self-expectations as the standard of judgment to determine whether or not an outcome is aversive.

Personal and normative self-standards can also be chronically accessible for particular individuals. If a person thinks of herself as being a deceitful person, she will not be upset at all by convincing a waiting fellow [p. 387 ↓] student that a boring task was actually interesting. She carries with her a chronically available self-standard and compares her behavior to that standard. Another person may think of herself as honest and carry that self-standard as a chronic measuring stick against which to judge her behavior. She would be in the throes of dissonance after agreeing to dupe a fellow student. For these two hypothetical students, their judgment is measured against a personal standard that overrides the effect of the social circumstance.

The predictions of the self-standard model have been supported in a number of studies (Weaver and Cooper, 2002; Stone, 1999; Stone and Cooper, 2003). When people compare their behavior to normative standards of judgment, then they assess consequences to be aversive in a manner similar to most people in the culture. We

would not expect dissonance to be moderated by their sense of self – for example, their level of self-esteem. By contrast, when ideographic dissonance is aroused by comparison to personal standards of judgment, then what is considered aversive varies by self-esteem. People with a high sense of self-esteem expect to make good and rational choices. They are upset when their choices lead to a consequence that is negative. When people with chronically low self-esteem make choices, they expect those choices to have negative results and are not upset by what other people would consider negative outcomes.

In a study reported by Stone (1999), participants were divided by median split into those with high and low self-esteem. They were asked to make choices between two attractive music albums. Half of the students were primed to make their personal standards accessible while the other half were primed to have their normative standards accessible. Following the decision of which album to keep, the participants rated the albums. The prediction was that normative-primed participants would experience dissonance whenever they made a difficult choice and would show the classic dissonance finding of raising their evaluation of the chosen album and reducing the attractiveness of the rejected album. Stone predicted that self-esteem would not enter into the findings because these students were measuring the consequences of their behavior against a normative standard of judgment. By contrast, self-esteem was expected to play a role when participants had been primed to use their personal standards. People with high self-esteem are far more likely than people with low self-esteem to believe that elements discrepant with their choice are aversive – but only because they are assessing those consequences against a personal standard of judgment. Stone found that when normative standards were primed, self-esteem made no difference and participants changed their attitudes toward the albums as predicted by dissonance. However, when personal standards were primed, participants with high self-esteem changed their attitudes far more than participants with low self-esteem.

Progress Report on a Classic Theory

Festinger thought that dissonance was a function of cognitive inconsistency. As I look back at more than a half century of theory and research on this now-classic theory, two facets of Festinger's genius are palpable from his writing. One facet was about form

and the other about substance. Festinger formed his theory around the great issues of the time. Learning was king of the psychology literature in the 1950s and Festinger adapted the Hullian drive concept for use in social psychology. In social psychology, learning and reinforcement concepts were the assumed girders of much of the research in persuasion and attitude change (Hovland et al., 1949, 1953). Festinger aligned his theory to make precisely the opposite predictions about persuasion than would have been predicted by learning, therefore creating instant controversy and research.

[p. 388 ↓] From a substantive perspective, Festinger thought that inconsistency among cognitions led to the uncomfortable arousal state he called dissonance. From what we now know, he was partially correct. "All theories are wrong," Festinger (1987) once wrote, "One asks, 'How much of the empirical realm can it handle and how must it be modified and changed as it matures?'" In the New Look and Self-Standards models, my colleagues and I tried to right the ship when it veered away from its path and to find new solutions to the *but only* dilemmas that taught us that the theory did not capture the totality of the dissonance phenomenon. To be clear, we were not the only investigators to notice that the theory needed additional concepts and perspectives to capture the richest range of phenomena and data. For example, Beauvois and Joule (1999), Harmon-Jones (1999), and Steele (1988) are among the creative scholars who have used alternate lenses to analyze the progress of dissonance through the decades. There is consensus that Festinger set us out on the path to understand how people view the 'fit' of their cognitions. There is consensus that one of his brilliant and ever-lasting insights was to allow us to consider all cognitions – whether mental representation of the world or mental representations of internal states – on the same grid and therefore subject to the rules of the dissonance process. There is consensus that his straightforward set of principles stimulated research in a way that was unprecedented in the field of social psychology.

I believe in the dissonance processes that we outlined in Cooper and Fazio (1984) and enhanced in Stone and Cooper (2001). As Festinger once taught us, however, our own version of the work will one day be proven wrong (only in part, I hope). Festinger wrote, "The only kind of theory that can be proposed and ever will be proposed that absolutely will remain inviolate for decades ... is a theory that is not testable. If a theory is at all testable, it will not remain unchanged. It has to change."

New Avenues of Dissonance Research

From Personal Dissonance to Vicarious Dissonance

People's selves are integral to the dissonance process. Recent theorizing has made clear that the self is both personal and social (Leary and Tangney, 2003). It is about one's own personal characteristics and simultaneously about one's interconnectedness with others and with social groups (e.g., Brewer and Gardner, 1996), yet prior research connecting the experience of cognitive dissonance to membership in social groups was scant during the formative period of dissonance research. Ironically, the first study of cognitive dissonance ever reported was the study of discontinued expectancies by members of a doomsday cult who believed that the world would end in a cataclysmic flood. Their reactions to the discontinued expectancy formed the basis of *When Prophecy Fails* (Festinger et al., 1956). However, it would take researchers several decades to systematically vary group membership and assess the impact of participants' acting as individuals compared with their acting as members of a small group (Cooper and Mackie, 1983; Zanna and Sande, 1987).

In the theory of vicarious dissonance (Cooper and Hogg, 2007), we went more to the heart of the meaning of group membership and considered its impact on dissonance. We considered the effect of one group member's counterattitudinal advocacy on the attitudes and behaviors of other members of one's group. Social identity theory was the vehicle that helped us to link the dissonant behavior of one group member with the attitudes of other members of the group. Because of the impact of social identity on members of social groups, we reasoned that dissonance aroused in one group member could cause other group members to experience dissonance vicariously and result in attitude change by the other members of the social group.

[p. 389 ↓] As we know from the important work in social identity theory (Tajfel, 1970) and social categorization theory (Turner and Hogg, 1987), people in groups forge a common identity. When thinking of themselves as group members, there is a tendency

toward depersonalization and intersubjectivity such that members assimilate toward the prototypical member of their group. The more strongly they feel about their group, the more they share in intersubjectivity and the more they take on the characteristics and emotions of fellow group members. Put simply, happiness or fear or sadness experienced by one member of a group spreads via intersubjectivity to other members (Mackie et al., 2007).

Michael Hogg and I wondered whether one group member's dissonance could spread to other group members in the same way. Suppose you are a member of a conservative antitax group and you observe a fellow group member make a public speech advocating an increase in the progressive income tax in order to support social programs. You know that the person voluntarily made the speech and that it was played before a potentially convincible audience. The situation has all of the ingredients to create cognitive dissonance in the speaker. But what about you, the witness? We reasoned that you will experience cognitive dissonance vicariously. Because of your common group membership, your view of your self is partly determined by your membership in the group to which you and the speaker belong. Your identity is wrapped up with your fellow group members and intersubjectivity makes you and the speaker fuse toward a common identity. The speaker's experience of discomfort will become your experience of discomfort. The speaker's attitude change will become your attitude change. The speaker's reduction of dissonance will become yours as well.

Norton et al. (2003) provided evidence for vicarious arousal of dissonance. A fictitious cover story created a rationale for a student to witness a fellow student agree to write a counterattitudinal message and to learn whether or not the student was a member of the participant's social group. At Princeton University, all entering undergraduate students are assigned at random to one of five residential colleges. Each student lives and eats in one of the colleges and each college has its own social and academic activities. The student's residential college served as the crux of the ingroup versus outgroup manipulation because each participant believed that he or she was witnessing an interaction with a student who happened to be a fellow resident of his or her residential college (ingroup member) or happened to be a resident of a different residential college (outgroup member).

The students arrived for a study of “linguistic subcultures” in groups of two, although each reported to a separate room, separated by two-way mirrors. We told the students that we were interested in how people in different residential colleges come to speak in slightly different ways, learning to use slightly different inflections or terms in their spoken behavior. For example, we know that people who live in the Midwest develop a slightly different pattern of speech than people who live in South Carolina or Massachusetts. The experimenter explained that the purpose of the students' participation in the current study was to see if these speech patterns occur in microcosms – that is, small groups within a larger context. We told the students that, in this study, we wanted to see if the speech patterns of students in the different residential colleges at Princeton University were different from one another and whether we could measure them.

We explained that one of the two students, selected at random, was going to deliver a speech on a given topic and the other student was going to listen carefully, and then respond to several questions about the speaker's speech patterns. Each participant was told that he or she was the one who had been randomly picked to rate the speech, while the student in the other room was assigned to give the speech. The procedure allowed us to make the student's residential college group salient and manipulate systematically whether [p. 390 ↓] the speaker's residential college group was the same (ingroup) or different (outgroup) from the participant's. The experimenter found a pretext to turn the lights on briefly, which allowed the participants to see that there truly was another student in the other room. The illumination was kept low so that the students' identities could not be accurately discerned. What students did not realize was that each of them had been assigned the role of listener. All information about what the alleged other student said or did was manipulated by instruction or audiotape.

The experimenter left the room, ostensibly to instruct the other participant about the speech he or she was to make. During the intervening period, participants filled out various measures, including measures of how much they liked and felt identified with their residential college on a scale developed by Hogg and his associates (Hogg et al., 1998). In a few minutes, the experimenter returned with a tape recording that included the completed speech and the experimenter's alleged conversation with the other student. On the tape, the experimenter explained that he was fortunate to be able to combine two studies into one. The dean's office had asked for a study trying to

assess student opinion about the possibility of raising tuition fees by a more than typical amount. The experimenter then asked the student to write a strong and forceful speech advocating a spike in tuition fees. He explained that this would be the speech that the other subject (i.e., the real participant) would rate for its linguistic features and that it would then be sent on to the dean's office. The experimenter also asked the alleged other student how he or she felt about raising tuition and the student responded, "Well ... I'd be against it."

The participants thus had a credible, albeit fabricated, story that allowed them to overhear an ingroup or an outgroup member make a counterattitudinal speech on a controversial topic. The tape recorder was stopped while the writer supposedly organized his or her thoughts, and then restarted for the participant to hear the alleged speech. The speech was a relatively brief exposition on how higher tuition rates could allow the university to hire more faculty staff, purchase more books for the library, and so forth. Before rating the speech for its linguistic properties, participants were asked about their own attitudes toward tuition increases at the university. This served as the dependent measure of our study.

The results of the study showed that observing a fellow group member behave in a counterattitudinal fashion caused the participant to change his or her attitude in the direction of the group member's counterattitudinal advocacy. As predicted by vicarious dissonance, this effect *only* occurred when the participant strongly identified with his or her group. In the absence of a strong affinity with one's group, observing an ingroup or an outgroup member did not affect participants' attitudes (see also Monin et al., 2004).

Dissonance, Vicarious Dissonance, and Culture

The concept of vicarious can help us unravel some of the cultural differences that have been identified in the expression of cognitive dissonance across cultures. Joan Miller (1984) was among the first to suggest that cultural differences may lead to different expressions of social psychological processes. She analyzed the difference between holistic and agentic cultures that broadly corresponded to Western European and

North American cultures on the one hand and Asian and Indian on the other. In agentic cultures, people see themselves as responsible for their own actions and decisions. They make personal attributions for events, viewing behavior as a window on their own traits and characteristics. People from holistic cultures view the self in relationship to others. They see situations and social roles as determining their behavior and view behavior as a means toward harmonious social relationships.

In a subsequent seminal paper, Markus and Kitayama (1991) expanded the analysis of cultural differences by drawing a distinction between collectivist and individualist [p. 391 ↓] cultures. Collectivist cultures are concerned with relationships among people. Social harmony is a key goal; attitudes and behavior primarily serve the goal of collective harmony. Individualist cultures are concerned with self-actualization. People's attitudes and behaviors are their own and expressing them truly and honestly is an important component of the self-actualization process.

By raising the question of whether different cultures have values that differentially affected the expression of attitudes, Markus and Kitayama opened an entirely new research direction for cognitive dissonance. They suggested that that dissonance is a uniquely Western or individualistic phenomenon. In individualistic cultures, people express opinions that are supposed to accurately reflect their judgment. They are supposed to say what they believe and believe what they say. Inconsistent cognitions do not fit an individualistic culture's notion of attitude expression. On the other hand, the expression of attitudes in collectivist cultures is only partly self-description but also constitutes expressions that affect the degree of harmony among people or groups. A member of a collectivist culture may not find it aversive to express attitudes that differ from their behaviors, but they would find it aversive if expressing opinions that disrupted interpersonal or intergroup harmony.

Although this chapter will not review the considerable research that has been conducted on dissonance in collectivist and individualist cultures during the last two decades, the conclusion of that research has revealed interesting and important aspects of the dissonance process itself. Heine and Lehman (1997) collected data in Canada comparing dissonance processes among Canadians of European descent and people of Japanese descent. Using a typical free-choice paradigm, Heine and Lehman found that, unlike the European Canadians, Japanese participants did not show the spreading

of alternatives effect that had been found numerous times in the social psychology literature.

Did this mean that cognitive dissonance is not experienced by people in collectivist cultures? Hoshino-Browne and colleagues (2005) reported a series of creative experiments in which they elucidated the impact of culture on cognitive dissonance. They showed that people from collectivist cultures showed dissonance reduction following a choice if they made that choice for a friend rather than for themselves. That is, when attitudes and behaviors were inconsistent within a social network of relationships, it produced dissonance. When the inconsistency had no social referent, it did not.

Research on culture and dissonance provides a window into the important social values that, when disrupted, create the aversive event that leads to cognitive dissonance. For collectivist cultures, the value is interpersonal harmony. When people behave in ways that upset the social order, it produces the aversive consequence that leads to cognitive dissonance. In Western cultures, when people act in ways that produce unwanted consequence to the individual actor, it leads to dissonance. Research by Kitayama and his colleagues has also shown that the same action that leads to dissonance arousal in individualists can also lead to dissonance among collectivists, when the presence of social others is subtly primed (Imada and Kitayama, 2010; Kitayama et al, 2004).

Vicarious dissonance research provides another perspective to consider the differences between individualist and collectivist cultures. Chong and Cooper (2007) reported a study using the induced compliance paradigm of dissonance. Students in South Korea wrote essays that could bring about an unwanted policy change at their university. Chong and Cooper found that Korean students did not change their attitudes following the counterattitudinal speech, even though they acted with free choice and their action had the potential of bringing about an unwanted policy. However, Korean students did change their attitudes when the situation was changed into a vicarious dissonance study. When they witnessed students from their group writing a counterattitudinal essay, they changed their own opinion just as the participants in Norton et al.'s (2003) study had done.

Vicarious dissonance is intrinsically a social phenomenon. It is an arousal that is [p. 392 ↓] experienced on behalf of someone else in an important social network. Taken together with the research of Hoshino-Browne et al. and Kitayama et al., we now understand that dissonance is aroused in both collectivist and individualistic cultures. For collectivists, more so than for individualists, a social consequence seems necessary for an act to be considered aversive and lead to the tension state of dissonance.

Dissonance in a Social World

Why Does it Matter?

One of the characteristics of cognitive dissonance is its ubiquity. When we make choices or suffer embarrassment or expend effort, we are in a dissonant state. It is difficult to go through a day without arousing dissonance. Attitudes ranging from consumer preferences (Menasco and Hawkins, 1978) to military service (Staw, 1974) have been viewed from a dissonance perspective. We have gained insight into why people feel passionately about social groups such as sororities and fraternities that they suffered to get into from the vantage point of cognitive dissonance. And what professor has not considered whether some students seem to love very difficult courses because of the effort justification intrinsic to cognitive dissonance?

To conclude this chapter, I will address two particular areas of social impact that have been addressed systematically with the principles of cognitive dissonance. The first is the potential of dissonance to effect change through psychotherapy and the second is the use of vicarious cognitive dissonance to induce positive changes in health behavior.

Dissonance as Psychotherapy

Can psychotherapy be considered an instance of cognitive dissonance? In the 1980s, Danny Axsom and I conducted a series of studies in which we showed how dissonance could induce people to change their attitudes and behaviors psychotherapeutically (Axsom, 1989; Axsom and Cooper, 1985; Cooper, 1980). We noted the parallel

between most psychotherapies and the principles of effort justification that Aronson and Mills had introduced in 1959. All therapies require effort. People engage in the effortful work freely, although the goal for which they are working is something for which they have some trepidation. It may be that they came to therapy to reduce their fear toward an object or their anxiety toward performing a particular behavior. Whatever the goal may be, prospective patients have ambivalence toward it, then engage in an effortful set of therapy sessions designed to overcome it.

We reasoned that the choice to engage in an effortful procedure was akin to the high choice conditions of an effort justification study. We decided to run studies in which people would attempt to reach a goal using a high degree of effort totally unrelated to any bona fide theory of psychotherapy. In one study, we invited people who wanted to reduce their fear of snakes to come to the laboratory where we measured how close they could come to our six-foot boa constrictor. They then participated in an effortful therapy that they believed was related to overcoming the fear. In truth, it contained a set of physical exercises designed to be difficult, embarrassing, and exhausting. We found that the participants were able to overcome their fear following this physical exercise therapy. Moreover, we also varied participants' choice to engage in the effortful therapy. Consistent with dissonance theory predictions, phobic participants who freely chose to engage in the effortful therapy overcame their fear significantly more than participants who were not given a choice (Cooper, 1980).

In a similar study, Axsom and I (Axsom and Cooper, 1985) put dissonance to use against the problem of obesity. People who were overweight and who had tried numerous weightloss programs volunteered for our experimental research. In a series of [p. 393 ↓] five sessions, we asked our experimental group to engage in tasks that required a great deal of cognitive effort. They made perceptual judgments, read tongue-twisters and recited nursery rhymes, backward, for an hour. A second group engaged only in a low degree of effort, spending their time making simple judgments and relaxing. Six months after the end of the sessions, the participants were weighed. As we had predicted from dissonance theory, the high effort group lost more weight than the low effort group (8.6 lbs versus 0.8 lbs respectively) and kept the weight off for a year.

We do not claim that all psychotherapy is cognitive dissonance but we do believe that cognitive dissonance is one of the active ingredients of most psychotherapies. With knowledge of the conditions that give rise to maximal dissonance, we should be able to design psychotherapies in a way that allows dissonance to be helpful for psychotherapeutic change. Whichever approach a particular therapist adopts, maximizing the impact of dissonance in the therapeutic program can only enhance the accomplishments of the therapy. Therapists are advised to focus patients' attention on the effortful nature of the patient's work. Moreover, therapists should emphasize the patient's personal responsibility for engaging in the therapeutic effort. If these elements are included in psychotherapy, then the arousal and motivation that accrue from the dissonance process will be put to productive use for the patient.

Vicarious Dissonance Can Lead to a Healthier Society

I believe that vicarious dissonance opens up a vast array of possibilities to put dissonance to use for a better world. We know that exposure to group members who are engaging in a dissonant act creates dissonance in other members (Monin et al. 2004; Norton et al, 2003). Consider the following scenario: A person observes a fellow group member advocate the use of risk-protective health behaviors; for example, eating a healthy diet, using condoms during sexual activity, refraining from smoking or applying sunscreen to protect against cancer. If the person observes the fellow group member admit to prior instances in which she did not practice what she preached, the elements for the vicarious experience of dissonance will be met. The vicarious dissonance should be resolved by the person's committing to more healthy future behaviors.

We (Fernandez et al., 2007) conducted such a study. Students at the University of Arizona were asked to listen to a speech made by a student that encouraged people to use sunblock as a preventative measure for skin cancer. They were led to believe that the speech had been made by a student at their university (ingroup condition) or by a student at a rival university (outgroup condition.) The speech was consistent with the participants' and the speaker's attitude toward sunscreen. It concluded, "No matter how

busy you think you are with work or school you can and should always wear sunscreen to reduce your risk of cancer."

Through the use of an appropriate cover story, participants heard the ingroup or outgroup speaker reminded that she did not use sunscreen herself *every time* she went outdoors. We predicted that the experience of vicarious dissonance would be aroused for students who heard an ingroup member admit to hypocrisy but not an outgroup member and the more identified a student was with her own group (University of Arizona), the more she would experience vicarious dissonance.

We predicted that the vicarious dissonance would lead to prohealth behaviors and prohealth attitude change on the part of the participants, and that is what we found. Women in the study changed their own attitudes to become more ardent in their opinion that sunscreen should be used on all occasions. Moreover, when given the opportunity to return a coupon for a free bottle of sunscreen, 74 percent of the women in the vicarious hypocrisy condition requested their bottle [p. 394 ↓] whereas only 54 percent of the women in the low vicarious hypocrisy (outgroup speaker) condition requested their sample.

There is good reason to believe that vicarious hypocrisy can be recruited by institutions from schools to the workplace to help its members live a healthier lifestyle and make healthier, less risky life decisions. Vicarious dissonance is a multiplier. A person who is induced to express vicarious hypocrisy creates his or her own change as an individual but may multiply that change throughout the groups to which he or she belongs. Schools, for example, can provide an opportunity for students to observe a fellow student speaking forcefully of her commitment to a healthy behavior regimen; for example, a commitment to a frequent exercise program. If the student acknowledges some occasions in which she failed to get to the gym, the conditions for vicarious dissonance will have been met. The dissonant cognitions of the speaker will spread to group members who will reduce their vicarious dissonance by adopting the exercise regimen that the single group member espoused. Similarly, a work place might bring group members together to witness a fellow worker advocate for healthy eating choices. By acknowledging his own dietary transgressions, fellow group members will experience vicarious dissonance, which they can reduce by making healthier food choices. The dissonance expressed by the single group member can spread through

the group, affecting all of the members of the group. We can further speculate that this will be especially true if the members feel strongly identified with their group.

Conclusion

I came upon the theory of cognitive dissonance when it was in its childhood. It was a precocious child, already having generated fans and enemies, proponents and critics. A half-century later, the theory continues to inspire. Although only a few of the most ardent skeptics still doubt the existence of dissonance, the precise mechanism continues to be elusive. My own joy from being part of the dissonance cohort comes from having helped move the theory to a new level of understanding. In my own thinking, the reliance on inconsistency gave way to an understanding of the role of responsibility-for-consequences that is the core of the New Look model and for the importance of the self-concept that is the foundation of the Self-Standard model.

As Festinger tried to teach us, anyone's perspective on dissonance process will ultimately prove inadequate as data continue to be collected. All theories are at least partly wrong and all theories must change. Nonetheless, the search for change is part of the science and part of the fun. In the half century since Festinger first brought dissonance to our attention, we have not only moved toward a deeper understanding of this ubiquitous process, but we have also seen the theory spawn new ideas and relationships. Such major theoretical approaches as Kunda's motivated reasoning (Kunda, 1990), Higgins' (1989) self-discrepancy and Tesser's (1988) self-evaluation maintenance are but a few examples of that search. There will be more. The theoretical stability of dissonance and the change it continues to inspire are the twin legacies of cognitive dissonance theory.

References

- Aronson, E. (1969) The theory of cognitive dissonance: A current perspective . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 7, 1–34. New York: Academic Press.

Aronson, E. (1992) The return of the repressed: dissonance theory makes a comeback . Psychological Inquiry , 3, 303–311. http://dx.doi.org/10.1207/s15327965pli0304_1

Aronson, E. and Carlsmith, J.M. (1962) Performance expectancy as a determinant of actual performance . Journal of Abnormal and Social Psychology , 65, 178–182. <http://dx.doi.org/10.1037/h0042291>

Aronson, E. and Carlsmith, J.M. (1963) The effect of the severity of threat on devaluation of forbidden behavior . Journal of Abnormal and Social Psychology , 66, 584–588. <http://dx.doi.org/10.1037/h0039901>

Aronson, E. and Mills, J. (1959) The effect of severity of initiation on liking for a group . Journal of Abnormal and Social Psychology , 59, 177–181. <http://dx.doi.org/10.1037/h0047195>

Axsom, D. (1989) Cognitive dissonance and behavior change in psychotherapy . Journal of Experimental Social Psychology , 25, 234–252. <http://dx.doi.org/10.1016/0022-1031%2889%2990021-8>

Axsom, D. and Cooper, J. (1985) Cognitive dissonance and psychotherapy: The role of effort justification in inducing weight loss . Journal of Experimental Social Psychology , 21, 149–160. <http://dx.doi.org/10.1016/0022-1031%2885%2990012-5>

Beauvois, J. and Joule, R.V. (1999) A radical point of view on dissonance theory . In E. Harmon-Jones, ed. and J. Mills (eds), Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology , pp. 43–70. Washington, DC: American Psychology Association. <http://dx.doi.org/10.1037/10318-003>

Brehm, J. (1956) Postdecision changes in the desirability of alternatives . Journal of Abnormal and Social Psychology , 52, 384–389. <http://dx.doi.org/10.1037/h0041006>

Brewer, M.B. and Gardner, W. (1996) Who is the ‘We’? Levels of collective identity and self representation . Journal of Personality and Social Psychology , 71, 83–93. <http://dx.doi.org/10.1037/0022-3514.71.1.83>

Carlsmith, J.M., Collins, B.E. and Helmreich, R.L. (1966) Studies in forces compliance: I. The effect of pressure for compliance on attitudes change produced by face to face role playing and anonymous essay writing . *Journal of Personality and Social Psychology* , 4, 1–13. <http://dx.doi.org/10.1037/h0023507>

Chong, J. and Cooper, J. (2007) Cognitive dissonance and vicarious dissonance in East Asia: Can I feel your discomfort but not my own? Poster presentation at a Meeting of Society for Personality and Social Psychology, Memphis, TN.

Cooper, J. (1980) Reducing fears and increasing assertiveness: The role of dissonance reduction . *Journal of Experimental Social Psychology* , 16, 199–213. <http://dx.doi.org/10.1016/0022-1031%2880%2990064-5>

Cooper, J. (1992) Dissonance and the return of the self-concept . *Psychological Inquiry* , 3, 320–323. http://dx.doi.org/10.1207/s15327965pli0304_5

Cooper, J. and Fazio, R.H. (1984) A new look at dissonance theory . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 17, 229–262. Hillsdale, NJ: Erlbaum.

Cooper, J. and Goethals, G.R. (1974) Unforeseen events and the elimination of cognitive dissonance . *Journal of Personality and Social Psychology* , 29, 441–445. <http://dx.doi.org/10.1037/h0036199>

Cooper, J. and Hogg, M.A. (2007) Feeling the anguish of others: A theory of vicarious dissonance . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 39, 359–403. San Diego, CA: Academic Press.

Cooper, J. and Mackie, D.M. (1983) Cognitive dissonance in an intergroup context . *Journal of Personality and Social Psychology* , 44, 536–544. <http://dx.doi.org/10.1037/0022-3514.44.3.536>

Cooper, J. and Worchel, S. (1970) The role of undesired consequences in the arousal and cognitive dissonance . *Journal of Personality and Social Psychology* , 312–320. <http://dx.doi.org/10.1037/h0028990>

Cooper, J., Zanna, M.P. and Goethals, G.R. (1974) Mistreatment of an esteemed other as a consequence affecting dissonance reduction . *Journal of Experimental Social Psychology* , 10, 224–233. <http://dx.doi.org/10.1016/0022-1031%2874%2990069-9>

Cooper, J., Zanna, M.P. and Taves, P. (1978) Arousal as a necessary condition for attitude change following induced compliance . *Journal of Personality and Social Psychology* , 36, 1101–1106. <http://dx.doi.org/10.1037/0022-3514.36.10.1101>

Croyle, R. and Cooper, J. (1983) Dissonance arousal: Physiological evidence . *Journal of Personality and Social Psychology* , 45, 782–791. <http://dx.doi.org/10.1037/0022-3514.45.4.782>

Davis, K.E. and Jones, E.E. (1960) Change in interpersonal perception as a means of reducing cognitive dissonance . *Journal of Abnormal and Social Psychology* , 61, 402–410. <http://dx.doi.org/10.1037/h0044214>

Dickerson, C.A, Thibodeau, R., Aronson, E. and Miller, D. (1992) Using cognitive dissonance to encourage water conservation . *Journal of Applied Social Psychology* , 22, 841–854. <http://dx.doi.org/10.1111/j.1559-1816.1992.tb00928.x>

Elliot, A.J. and Devine, P.G. (1994) On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort . *Journal of Personality and Social Psychology* , 67, 382–394. <http://dx.doi.org/10.1037/0022-3514.67.3.382>

Fazio, R.H., Zanna, M.P. and Cooper, J. (1977) Dissonance and self-perception: An integrative view of each theory's proper domain of application . *Journal of Experimental Social Psychology* , 13, 464–479. <http://dx.doi.org/10.1016/0022-1031%2877%2990031-2>

Fernandez, N., Stone, J., Cascio, E., Cooper, J. and Hogg, M.A. (2007) Vicarious hypocrisy: The use of attitude bolstering to reduce dissonance after exposure to a hypocritical ingroup member . Paper presented at the Meeting of the Society for Personality and Social Psychology, Memphis, TN.

Festinger, L. (1950) Informal social communication . *Psychological Review* , 57, 271–282. <http://dx.doi.org/10.1037/h0056932>

Festinger, L. (1954) A theory of social comparison processes . *Human Relations* , 7, 117–140. <http://dx.doi.org/10.1177/001872675400700202>

Festinger, L. (1957) A Theory of Cognitive Dissonance . Evanston, IL: Row, Peterson.

Festinger, L. (1987) Reflections on cognitive dissonance: Thirty years later . Paper presented at the 95th Annual Convention of the American Psychological Association. August, 1987, New York.

Festinger, L. and Carlsmith, J.M. (1959) Cognitive consequences of forced compliance . *Journal of Abnormal and Social Psychology* , 58, 203–210. <http://dx.doi.org/10.1037/h0041593>

Festinger, L., Riecken, H.W. and Schachter, S. (1956) When Prophecy Fails . Minneapolis, MN: University of Minnesota Press. <http://dx.doi.org/10.1037/10030-000>

Freedman, J.L. (1965) Long-term behavioral effects of cognitive dissonance . *Journal of Experimental Social Psychology* , 1, 145–155. <http://dx.doi.org/10.1016/0022-1031%2865%2990042-9>

Fried, C.B. and Aronson, E. (1995). Hypocrisy, misattribution, and dissonance reduction . *Personality and Social Psychology Bulletin* , 21, 925–933. <http://dx.doi.org/10.1177/0146167295219007>

Goethals, G.R. and Cooper, J. (1972) Role of intention and postbehavioral consequences in the arousal of cognitive dissonance . *Journal of Personality and Social Psychology* , 23, 292–301. <http://dx.doi.org/10.1037/h0033123>

Goethals, G.R., Cooper, J. and Naficy, A. (1979) Role of foreseen, foreseeable, and unforeseeable behavioral consequences in the arousal of cognitive dissonance . *Journal of Personality and Social Psychology* , 37, 1179–1185. <http://dx.doi.org/10.1037/0022-3514.37.7.1179>

Harmon-Jones, E. (1999) Toward an understanding of the motivation underlying dissonance effects: Is the production of aversive consequences necessary? In E. Harmon-Jones, ed. and J. Mills (eds), *Cognitive Dissonance: Progress on a Pivotal*

Theory in Social Psychology , pp. 71–99. Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10318-004>

Harmon-Jones, E., ed. and Mills, J. (eds) (1999) Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology , Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10318-000>

Harmon-Jones, E., Brehm, J.W., Greenberg, J., Simon, L. and Nelson, D.E. (1996) Evidence that the production of aversive consequences is not necessary to create cognitive dissonance . Journal of Personality and Social Psychology , 70, 5–16. <http://dx.doi.org/10.1037/0022-3514.70.1.5>

Heider, F. (1946) Attitudes and cognitive organization . The Journal of Psychology , 21, 107–112. <http://dx.doi.org/10.1080/00223980.1946.9917275>

Heine, S.J. and Lehman, D.R. (1997) Culture, dissonance, and self-affirmation . Personality and Social Psychology Bulletin , 23, 389–400. <http://dx.doi.org/10.1177/0146167297234005>

Higgins, E.T. (1989) Self-discrepancy theory: what patterns of self-beliefs cause people to suffer? In L. Berkowitz (ed.) Advances in Experimental Social Psychology , 22, 93–1360. San Diego, CA: Academic Press.

Hogg, M.A., Haines, S.C. and Mason, I. (1998) Identification and leadership in small groups: Salience, frame of reference, and leader stereotypicality effects on leader evaluations . Journal of Personality and Social Psychology , 75, 1248–1263. <http://dx.doi.org/10.1037/0022-3514.75.5.1248>

Hoshino-Browne, E., Zanna, A.S., Spencer, S.J., Zanna, M.P., Kitayama, S. and Lackenbauer, S. (2005) On the cultural guises of cognitive dissonance: The case of Easterners and Westerners . Journal of Personality and Social Psychology , 89, 294–310. <http://dx.doi.org/10.1037/0022-3514.89.3.294>

Hovland, C.I., Lumsdaine, A.A. and Sheffield, F.D. (1949) Experiments on Mass Communication (Studies in Social Psychology in World War II, Vol. 3 . Princeton, NJ: Princeton University Press.

Hovland, C.I., Janis, I.L. and Kelley, H.H. (1953) *Communication and Persuasion* . New Haven: Yale University Press.

Kunda, Z. (1990) The case for motivated reasoning . *Psychological Bulletin* , 8, 480–498. <http://dx.doi.org/10.1037/0033-2909.108.3.480>

Imada, T. and Kitayama, S. (2010) Social eyes and choice justification: Culture and dissonance revisited . *Social Cognition* , 28, 589–608. <http://dx.doi.org/10.1521/soco.2010.28.5.589>

Kitayama, S., Snibbe, A.C., Markus, H.R. and Suzuki, T. (2004) Is there any “free” choice? Self and dissonance in two cultures . *Psychological Science* , 15, 527–533. <http://dx.doi.org/10.1111/j.0956-7976.2004.00714.x>

Leary, M.R. and Tangney, J.P. (2003) *Handbook of Self and Identity* . New York: Guilford Press.

Lewin, K. (1951) *Field Theory in Social Psychology* . New York: Harper.

Under, D.E., Cooper, J. and Jones, E.E. (1967) Decision freedom as a determinant of the role of incentive magnitude in attitude change . *Journal of Personality and Social Psychology* , 6, 245–254. <http://dx.doi.org/10.1037/h0021220>

Losch, M.E. and Cacioppo, J.T. (1990) Cognitive dissonance may enhance sympathetic tonus, but attitudes are changed to reduce negative affect rather than arousal . *Journal of Experimental Social Psychology* , 26, 289–304. <http://dx.doi.org/10.1016/0022-1031%2890%2990040-S>

Mackie, D.M., Maitner, A.T. and Smith, E.R. (2007) Intergroup emotions theory . In T.D. Nelson (ed.), *Handbook of Prejudice, Stereotyping, and Discrimination* , pp. 285–307. Mahwah, NJ: Erlbaum.

Markus, H. and Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation . *Psychological Review* , 98, 224–253. <http://dx.doi.org/10.1037/0033-295X.98.2.224>

Menasco, M.B. and Hawkins, D.L. (1978) A field test of the relationship between cognitive dissonance and state anxiety . *Journal of Marketing Research* , 15, 650–655. <http://dx.doi.org/10.2307/3150639>

Miller, J.G. (1984) Culture and the development of everyday social explanation . *Journal of Personality and Social Psychology* , 46, 961–978. <http://dx.doi.org/10.1037/0022-3514.46.5.961>

Monin, B., Norton, M.I., Cooper, J. and Hogg, M.A. (2004) Reacting to an assumed situation vs. conforming to an assumed reaction: The role of perceived speaker attitude in vicarious dissonance . *Group Processes and Intergroup Relations* , 7, 207–220. <http://dx.doi.org/10.1177/1368430204046108>

Newcomb, T.M. (1956) The prediction of interpersonal attraction . *American Psychologist* , 11, 575–586. <http://dx.doi.org/10.1037/h0046141>

Norton, M.I., Monin, B., Cooper, J. and Hogg, M.A. (2003) Vicarious dissonance: Attitude change from the inconsistency of others , *Journal of Personality and Social Psychology* , 85, 47–62. <http://dx.doi.org/10.1037/0022-3514.85.1.47>

Pallak, M.S. and Pittman, T.S. (1972) General motivational effects of dissonance arousal . *Journal of Personality and Social Psychology* , 21, 349–358. <http://dx.doi.org/10.1037/h0032266>

Rosenberg, M. (1965) Society and the Adolescent Self-image . Princeton, NJ: Princeton University Press.

Scher, S. and Cooper, J. (1989) The motivational basis of dissonance: The singular role of behavioral consequences . *Journal of Personality and Social Psychology* , 56, 899–906. <http://dx.doi.org/10.1037/0022-3514.56.6.899>

Skinner, M.L. (1953) Science and Human Behavior . Oxford: Macmillan.

Staw, B.M. (1974) Attitudinal and behavioral consequences of changing a major organizational reward: A natural field experiment . *Journal of Personality and Social Psychology* , 29, 742–751. <http://dx.doi.org/10.1037/h0036292>

Steele, C.M. (1988) The psychology of self-affirmation: sustaining the integrity of the self . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 21, 261–302. San Diego, CA: Academic Press.

Stone, J. (1999) What exactly have I done? The role of self-attribute accessibility in dissonance . In E. Harmon-Jones, ed. and J. Mills (eds), *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology* , pp. 175–200. Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10318-008>

Stone, J. and Cooper, J. (2001) A self-standards model of cognitive dissonance . *Journal of Experimental Social Psychology* , 37, 228–243. <http://dx.doi.org/10.1006/jesp.2000.1446>

StoneJ. and Cooper, J. (2003) The effect of self-attribute relevance on how self-esteem moderates attitude change in dissonance processes , *Journal of Experimental Social Psychology* , 39, 508–515. <http://dx.doi.org/10.1016/S0022-1031%2803%2900018-0>

Stone, J., Aronson, E., Crain, A.L., Winslow, M.P. and Fried, C.B. (1994) Inducing hypocrisy as a means of encouraging young adults to use condoms , *Personality and Social Psychology Bulletin* , 20, 116–128. <http://dx.doi.org/10.1177/0146167294201012>

Tajfel, H. (1970) Experiments in intergroup discrimination . *Scientific American* , 223, 96–102. <http://dx.doi.org/10.1038/scientificamerican1170-96>

Tesser, A. (1988) Toward a self-evaluation maintenance model of social behavior . In L. Berskowitz (ed.) *Advances in Experimental Social Psychology* , 21, 181–227. New York: Academic Press.

Thibodeau, R. and Aronson, E. (1992) Taking a closer look: Reasserting the role of the self-concept in dissonance theory . *Personality and Social Psychology Bulletin* , 18, 591–602. <http://dx.doi.org/10.1177/0146167292185010>

Turner, J.C. and Hogg, M.A. (1987) *Rediscovering the Social Group* . Oxford: Blackwell.

Waterman, C.K. and Katkin, E.S. (1967) Energizing (dynamogenic) effect of cognitive dissonance on task performance . *Journal of Personality and Social Psychology* , 6, 126–131. <http://dx.doi.org/10.1037/h0024573>

Weaver, K.D. and Cooper, J. (2002) Self-standard accessibility and cognitive dissonance reduction . Paper presented at the Meeting of the Society for Personality and Social Psychology. Palm Springs, CA, January, 2002.

Zanna, M.P. and Cooper, J. (1974) Dissonance and the pill: An attribution approach to studying the arousal properties of dissonance . *Journal of Personality and Social Psychology* , 29, 703–709. <http://dx.doi.org/10.1037/h0036651>

Zanna, M.P. and Sande, G.N. (1987) The effects of collective actions on the attitudes of individual group members: A dissonance analysis . In M.P. Zanna, ed. , J.M. Olson, ed. , and C.P. Herman (eds), *The Ontario Symposium: Vol. 5. Social Influence* , pp. 151–163. Hillsdale, NJ: Lawrence Erlbaum Associates.

<http://dx.doi.org/10.4135/9781446249215.n19>

Handbook of Theories of Social Psychology: Volume 1

Terror Management Theory

Contributors: Jeff Greenberg & Jamie Arndt

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Terror Management Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n20>

Print pages: 398-416

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n20>

[p. 398 ↓]

Chapter 19: Terror Management Theory

Jeff Greenberg, ed. and Jamie Arndt, ed.

Abstract

Terror management theory was developed to explain the motivational underpinnings of phenomena such as self-esteem defense and prejudice. The theory is rooted in a long tradition of thought regarding human awareness of death and its role in psychological functioning. The theory posits that to manage the potential for terror engendered by the awareness of mortality, humans sustain faith in worldviews which provide a sense that they are significant beings in an enduring, meaningful world rather than mere material animals fated only to obliteration upon death. The theory is supported by a wide range of studies showing that self-esteem and worldviews provide protection against anxiety and death-related cognition, reminders of mortality instigate worldview bolstering and self-esteem striving, and threats to the worldview and self-esteem increase the accessibility of death-related thought. The research has also led to a dual defense model of responses to conscious and unconscious death thoughts. We then focus on two of many topics informed by the theory: attitudes and behavior regarding physical health, and political preferences and intergroup conflict. We then consider factors that mitigate destructive forms of terror management. Finally, we briefly summarize the contribution of terror management work so far and where it's heading.

Terror Management Theory

Terror management theory was born out of a fundamental dissatisfaction with the field of social psychology circa 1980 shared by three graduate students at the University of Kansas: Sheldon Solomon, Tom Pyszczynski, and the senior author of this chapter. Firmly in the grip of social cognition, the field portrayed people as dispassionate information processors guided by schemas and heuristics, operating in a historical, cultural, motivational, and affective vacuum. Raised in working class families characterized by joy and anger, sibling rivalry and love, passion and sarcasm, in neighborhoods centered around churches and synagogues, bars and ballparks, full of people driven by regional, ethnic, and occupational pride and conflict, social psychology at the time struck us as describing some affectless androids the three of us had never met.

In viewing people as highly motivated creatures driven by their needs to protect their self-esteem and assert the superiority of [p. 399 ↓] their own group over other groups, we were like young salmon swimming upstream against the cognitive revolution. We focused our research on how motivation influenced people's perceptions of themselves and others, particularly in the context of defending self-esteem. We contributed to research on self-serving biases and self-handicapping, which showed that people bias their cognitions to protect their self-esteem. But we completed our graduate studies believing that we had no idea what drove people toward their prides and prejudices.

Spread around the country by the whims of the difficult job market (some things never change), the three of us began searching outside the discipline's journals for answers. I recall getting a call from Sheldon, back in 1982, in which he proclaimed that he had found a guy with the answers. It was a deceased cultural anthropologist named Ernest Becker, and the answers were in his book, *The Denial of Death*, which had won the 1973 Pulitzer Prize for general nonfiction. I quickly read the book. I found it frightening, brutal, and revelatory. Building on a wide array of sources, and taking an existential psychoanalytic perspective, the book seemed to explain all the human trends we had observed growing up and which were so well documented by social psychological research: conformity, obedience, self-serving biases, aggression, and prejudice. It explained everything from the rise of Nazism in Germany to the complexities of sex.

Sheldon, Tom, and I began discussing the book's many insights, as well as those provided by his earlier book, *The Birth and Death of Meaning* (1971), and his last book, *Escape from Evil* (1975). We began using Becker's ideas to teach social psychology in a more integrated way, as if the creature being discussed when we covered self-esteem was the same one we discussed when we covered social influence, aggression, prejudice, and close relationships. We were invited by Roy Baumeister to participate in a symposium entitled "Public and Private Self" at the October 1984 Society of Experimental Social Psychology (SESP) conference in Snowbird, Utah, and to contribute a chapter to an associated edited volume. We decided at SESP to introduce the core of Becker's analysis to our colleagues in social psychology, though Tom couldn't make it. On a snowy morn in front of an unlit fireplace in our room at the Snowbird Ski Lodge, Sheldon and I rather excitedly hacked out a simple summary of the theory, which we there dubbed "terror management theory" (TMT).

The symposium room was moderately full as Sheldon started his presentation, but as he began discussing Marx, Kierkegaard, Freud, and Otto Rank, a good half of the audience jostled their way out of the room. From the back of the room, I spotted a few remaining dignitaries such as John Darley, Ned Jones, who was visibly shaking his head throughout the talk, and our graduate school advisor and one fan, Jack Brehm. At talk's end, instead of the high fives we had envisioned, we received stunned silence, shock, and dismay. I was comforted at least that the practice of stoning had gone out of fashion.

Undeterred by the chilly initial reception, we forged ahead with the chapter, which became the first written presentation of TMT (Greenberg et al., 1986). Simultaneously, we worked on a paper more fully presenting the theory and explaining its potential value as a broad explanatory framework. We were hoping such a paper would be well received at *American Psychologist*, a publication which clamors for broad, integrative ideas. The paper was rejected unceremoniously with two reviews; one a single paragraph, the other a single sentence: "I have no doubt that this paper would be of no interest to any psychologist, living or dead."

Socialized into our profession very well, we did not accept these concise but argument-free bases for rejection. After a year or so of back and forth, editor Leonard Eron explained that the ideas may have had some validity but would not gain currency until

backed by empirical research. Actually, up to that moment, we had not thought of TMT as something to study empirically. But it struck [p. 400 ↓] us – duh – that this was precisely what we had been trained to do – to derive testable hypotheses from theories, and then test them.

We believe that TMT explains some fundamental things about human beings and their social behavior. But TMT was not like your standard variety social psychological theory. Most theories in social psychology are mini-theories focused on a specific set of processes pertinent to a particular topic within the field: stereotype threat (prejudice); the elaboration likelihood model (persuasion); the culture of honor (aggression); self-verification theory (self). TMT is about the role of the unconscious fear of death in just about everything we humans do. But soon we were able to derive hypotheses from this broad existential psychodynamic theory and strategies for testing them in collaboration with our graduate students.

Realizing resistance was going to be strong, we put together six studies, a larger set than was customary at that time, before submitting to the *Journal of Personality and Social Psychology*. The paper was accepted (Rosenblatt et al., 1989); the reviews were along the lines of, “This can’t be right, I don’t like it, but I can’t explain their findings, so take it.” We thought at the time, “Fair enough.”

Since then, the empirical support for TMT has blossomed into a literature consisting of over 400 studies and counting, conducted in 16 countries. This body of work, which has included numerous theoretical expansions and refinements, has benefited greatly from the second (and now third) generation of TMT researchers, former students of the original trio including the junior author of this chapter. In the last decade, researchers from independent labs throughout the world have also made valuable contributions to this continually expanding literature. We believe that the growth of this area of study reflects the generative value of a broad theory that integrates widely varied human endeavors by exploring deeply rooted forces that drives humans to behave the way they do. Indeed, the junior author of this chapter became enthralled as an undergraduate at Skidmore College with the broad existential perspective TMT offered and the promise of subjecting such ideas to empirical scrutiny. In due course, we will provide a brief overview of the literature that has developed over the past 20 years with an emphasis

on recent directions pertinent to contemporary concerns. But first we should step back to acknowledge the rather extensive roots of the theory.

The Distal and Proximal Roots of TMT

Although met with surprise and skepticism when introduced to social psychologists in the 1980s, a decent case can be made that TMT is an ancient theory that can be traced back to one of the first narrative texts from around 3000 BC, the Sumerian Epic of Gilgamesh. A central theme of this story, which influenced the major religions that later emerged out of the Middle East, is the main character's deep concerns about death and his consequent search for immortality. Gilgamesh, devastated by the death of his friend Enkidu, realizes that he too will die. He roamed the desert, lamenting: "How can I rest, how can I be at peace? Despair is in my heart. What my brother is now, that shall I be when I am dead ... I am afraid of death..." He then embarked on a journey to find immortality.

From that early time on, the notion that we humans fear death and urgently desire to deny or transcend it in some way has been a prominent theme in literature, religious writings, and philosophy. Indeed, Schopenhauer declared death the muse of philosophy; we couldn't even begin to review its role in philosophical thought here. But we should note that the first person to put the basic points of TMT together seems to have been the famed Greek historian Thucydides.

Around 400 BC, Thucydides was focused on understanding the vicious intergroup [p. 401 ↓] conflicts that plagued Ancient Greece. He proposed that the fear of inevitable death drove people to seek immortality in three ways: through heroic, noble actions restoring justice, which qualified them for a divinely awarded afterlife; through memories of their heroic deeds; and through identification with death transcending group identifications. As Ahrendorf (2000: 591) put it: "Thucydides contends that they will seek to escape by somehow overcoming their mortal condition, by living on after their death – either through their city, or through their glory, or in an afterlife – and by winning the gods' favor through the vehement affirmation of their own nobility or piety, or justice." Thucydides also noted that the increasing salience of mortality once conflicts begin intensifies the desire to heroically vanquish the enemy.

Jumping forward about 2000 years to the modern English literary tradition, poets from Shakespeare to Wordsworth, Keats, and Shelley, to Dickinson and Emerson, among hundreds of others, have acknowledged the role of the fear of death and desire to escape it in the human psyche. Similarly, novelists such as Swift, Dickens, Dostoevsky, and Tolstoy down to twentieth-century writers such as James Baldwin, Don DeLillo, James Joyce, Phillip Roth, Milan Kundera, and Kurt Vonnegut have explored how the fear of death drives various forms of human behavior. Here, Baldwin concisely captures the gist of TMT:

Life is tragic simply because the earth turns and the sun inexorably rises and sets, and one day, for each of us, the sun will go down for the last, last time. Perhaps the whole root of our trouble, the human trouble, is that we will sacrifice all the beauty of our lives, will imprison ourselves in totems, taboos, crosses, blood sacrifices, steeples, mosques, races, armies, flags, nations, in order to deny the fact of death, which is the only fact we have (James Baldwin, *The Fire Next Time*, 1963).

Awareness of the role of death in the human psyche has not been limited to the great philosophers, poets, and novelists, but can also be found extensively in visual art (e.g., Van Gogh, Klimt), music (e.g., Schubert and Mahler), and film (e.g., Woody Allen and Ingmar Bergman). And it is not limited to representatives of "high culture;" indeed, it seems increasingly prevalent in contemporary pop culture also. As we've been writing this chapter in early 2009, two films, David Fincher's Fitzgerald-inspired *The Curious Case of Benjamin Button*, and Frank Miller's comic-book-inspired *The Spirit*, as well as the second 2009 episode of the sitcom *Scrubs* have focused squarely on the psychological problem of mortality. Recognition of this problem in so many of the products of Western culture makes it all the more interesting that in the 1980s social psychology texts made no mention of death and TMT was treated by most people in the field as outlandish and irrelevant to understanding human social behavior.

As noted earlier, TMT was formalized out of three of Becker's books, *The Birth and Death of Meaning*, *The Denial of Death*, and *Escape from Evil*. These books combine insights from anthropology, evolutionary biology, philosophy, psychoanalysis, and sociology. We urge anyone interested in understanding human behavior to read them.

Around the time Becker was putting his synthesis together, evolutionary philosopher Susanne Langer and psychoanalytic historian Robert Jay Lifton were coming to the same view, and a bit later in 1980, so did existential psychotherapist Irvin Yalom. The primary influences on Becker's analysis were Kierkegaard, William James, Freud, Gregory Zilboorg, Erving Goffman, Norman Brown, and especially Otto Rank. Within psychology, Rank, Freud's disciple and an impressive interdisciplinary scholar back in the 1930s, first granted the fear of death and the urge toward immortality a central role in human culture and social behavior. By systematizing these existential psychodynamic ideas into a coherent explanatory framework, TMT integrates much of what we know about human behavior and provides a fertile basis of generating a wealth of testable hypotheses. The TMT research in turn has inspired useful expansions and refinements of the theory.

[p. 402 ↓]

The Basics of TMT

TMT starts with two simple observations. The first is that humans are animals with many systems designed to keep them alive, including a fight-or-flight response to impending threats to their continued existence. The second is that the cognitive abilities of humans have led to the awareness that death is inevitable and could come at any time due to a host of potential causes. The theory posits that this knowledge, in an animal built to avert death, creates an ever-present potential for intense anxiety, or terror, which must be managed continuously. The terror is managed by sustaining faith in a view of the world and oneself that denies the precarious and transient nature of one's own existence.

Cultural worldviews have been shaped to serve this function from ancient times to now. Presumably there was a point in time when the cerebral cortex of our species became sufficiently developed to provide self-awareness and the ability to think in terms of past, present, and future. These largely adaptive cognitive capacities also led to the awareness of mortality. Although fear in response to imminent threat is often adaptive, constant anxiety regarding the fragility of continued existence and its inevitable end would not be.

At this point, our ancestors constructed and gravitated toward shared conceptions of reality that most effectively quelled the potential fear that knowledge of our vulnerability and mortality engendered. These cultural worldviews imbue external reality with order, stability, meaning, and purpose, and offer ways in which people can believe they will endure beyond death either literally through an everlasting soul, symbolically through a death transcending identity, or for most worldviews, in both ways.

At the most basic level, all cultural world-views allow people to live out their lives largely within a conception of reality in which they view themselves as symbolic or spiritual beings who exist in a meaningful world, rather than as mere transient animals fated only to obliteration upon death. Indeed, the content of consciousness is structured by the cultural worldview in which a given individual is raised. We think in terms of names, dates, months, days of the week, hours, minutes, social roles, and categories. But these are ultimately elaborate window dressing that lends artificial and largely arbitrary structure to an ongoing subjective experience of unique sensations and perceptions that begins at birth and ends at death. Questions like “Who are you?” “Where are you?” “What time is it?” can only be answered with reference to culturally created constructs.

How do people become imbedded in a cultural worldview that provides them with their fundamental psychological security? Developmentally, the human newborn is the most helpless and dependent of all living creatures. It is also a highly distressed creature because, as Rank (1932 /1989) noted, it was conceived and developed attached to its mother and was suddenly, violently separated from its warm home, the womb, at birth. From exposure to the first words that flow out of their parents' mouths, the infant begins to be socialized into the prevailing cultural worldview. In early childhood, the only basis of psychological security for these helpless creatures is the care and love of the parents. The parents provide knowledge, comfort, nutrition, and protection.

As the parents begin to require particular actions from the child to sustain their love and approval, the child learns that when she does the right things, all is good in the world but when she does the wrong things, all hell could break lose. To sustain the psychological security of the parental love, the child internalizes the parental values of good and bad, and tries to live up to the good, despite the infringement on the child's natural desires and urges this entails. For the good little five-year-old sitting on mom's lap, all is right in the world. For the bad five-year-old who just punctured one of dad's

new stereo speakers with a flying action figure, the world has just become a scary place. In this way, the child comes to associate being good and valued [p. 403 ↓] with feeling secure, and being bad and worthless with dread. Self-esteem, that feeling of being valuable, thus comes to serve an anxiety-buffering function.

The plot thickens as the child's cognitive development leads her to realize that the things she fears – the dark, monsters, ghosts, big dogs – threaten her very existence and that the parents are limited and can't always protect them from everything. Concomitant with these dawning realizations, the child's basis of security gradually shifts to larger more powerful sources, deities, presidents, and the culture in general. This transference process is facilitated by the fact that the parents have been inculcating the child into the larger worldview, including the parent's own bases of psychological security, throughout the child's upbringing.

So over childhood, psychological security becomes predicated on being a good Christian, American, and so forth, on being valuable within the context of the individual's internalized version of the cultural world-view. This sense of enduring significance is not only linked implicitly with safety and security but explicitly with the literal immortality of heaven and the symbolic immortality of being part of death-transcending entities such as the family line and the nation, and lasting cultural achievements in science, politics, and the arts. In these ways, the significant person in a society can feel like he or she is an eternal part of an enduring reality.

Summary of TMT and its Basic Implications

In sum, the simple formula for effective terror management is faith in a meaning providing cultural worldview and the belief that one is a valuable contributor to that meaningful world (the TMT conceptualization of self-esteem). We initially derived two basic implications from the theory (Solomon et al., 1991). First, self-esteem and the worldview upon which it is predicated serve a critical anxiety-buffering function. So people strive for and defend self-esteem in the service of psychological security and these efforts depend on the specific cultural worldview to which the individual subscribes. Second, because these constructs are ultimately fragile social constructions, people will react negatively to anyone or anything that undermines faith in

their worldview or their self-worth. We believe that this provides a very basic insight into prejudice and intergroup conflict. People who criticize one's own worldview, or simply adhere to a worldview very different from one's own, call into question the validity of one's own basis of psychological security.

TMT therefore posited that such different people are inherently threatening and are reacted to with four defenses. The first, most widespread, is derogation. If these others are ignorant or evil, then their alternative beliefs can be dismissed. The second is assimilation. If these people are wrong, then I can help them see the light, which will make me all the more certain my worldview is the right one. Missionary activity provides a vivid example of this strategy. The third is accommodation, which involves incorporating appealing aspects of the alternative worldview into one's own. In this way, one can preserve faith in the central components of one's own worldview. Rock and rap music initially expressing anger and challenging mainstream American culture lose their threatening nature when converted to muzak or used to sell fast food burgers. Finally there is annihilation. If one's worldview is right and the others' is wrong, then one's own should prevail.

The Core Tests of Hypotheses Derived from TMT

TMT is consistent with a wide range of evidence from anthropology, archeology, and history, but our challenge was to deduce predictions that put the theory to the test. The central TMT hypotheses are based on its account of the psychological function of cultural worldviews and self-esteem. If these structures provide [p. 404 ↓] protection against death-related fears, then: (a) reminding people of their mortality should increase bolstering of the worldview and striving for self-worth; (b) bolstering these structures should reduce anxiety in response to threat and defensive reactions to reminders of death; and (c) threatening the structures should arouse anxiety and bring death-related concerns closer to consciousness.

Hundreds of studies have supported these hypotheses. First, let's consider hypothesis (a). Reminding people of their mortality (mortality salience or MS) increases positive

reactions to people who validate aspects of participant's worldviews and negative reactions to people who challenge aspects of their worldview or espouse a different one (see Greenberg et al., 2008 for a review). For example, MS leads to negative evaluations of Jews by Christians, of a critic of the US by Americans, and of foreign products by Germans. MS even leads to increased aggression against a critic of one's preferred political orientation, whether liberal or conservative. On the other hand, MS increases positive reactions to heroes and celebrities, one's national soccer team, and members of one's own religious group, and leads to more donations to valued charities. MS also increases striving for self-worth (see Greenberg et al., 2008). In people who value such things, MS increases bold driving, displays of physical strength, proenvironmental intentions, focus on one's appearance, interest in self-esteem enhancing dating partners, and ratings of the importance of and desire for fame and wealth.

A substantial number of studies have also supported hypothesis (b). Bolstering self-esteem reduces anxiety in response to threat (Greenberg et al., 1992b) and defensive reactions and death-thought accessibility following reminders of mortality (Harmon-Jones et al., 1997). Similarly, bolstering or defending one's worldview also reduces defensive reactions and death-thought accessibility following reminders of mortality (e.g., Arndt et al., 1997).

Recent research has also supported Hypothesis (c). Reminders that humans are animals, criticism of participants' world-views, and threats to participants' self-esteem all increase the accessibility of death-related thoughts and not the accessibility of other negative thoughts (e.g., Friedman and Rhoades, 2007; Schimel et al., 2007).

This body of research has operationalized reminding people of their mortality in a variety of ways and has extensively compared reminders of mortality with reminders of other negative thoughts (see Greenberg et al., 2008). People have been reminded of death via two simple questionnaire items, writing one sentence about death, fear of death scales, proximity to cemeteries and funeral homes, and subliminal death primes. Thinking about mortality has been compared to thinking about being paralyzed, failing, feeling uncertain, general anxieties, giving a speech in public, intense pain, dental pain, disease, unpredictable, intense bouts of pain, life being meaningless, being socially excluded, upcoming exams, unexpected events, and worries after college.

Although some studies have found similar effects of other aversive thoughts in circumscribed contexts (e.g., McGregor et al., 2001; van den Bos, 2001) the vast majority have found quite different effects for reminders of mortality. TMT theorists (e.g., Greenberg et al., 2008) have argued that this is because death is the only inevitable future event, it is what human biological systems, including the fight or flight system, are most focused on avoiding, and it threatens to undermine all human desires, whether for belonging, cognition, control, or growth. One possibility is that when these other threats lead to defenses similar to those instigated by MS, they may undermine structures that serve terror management, thereby bringing thoughts of death closer to focal awareness. For example, Landau et al. (2004a) found that threatening individuals' belief in a just world led to elevated death-thought accessibility. These other aversive experiences may also, under some conditions, be threatening in their own right. Given the clear evidence that thoughts of mortality so often arouse different responses [p. 405 ↓] than these other aversive cognitions, an important direction for future inquiry research is understanding when psychological defenses are serving terror management and when they are serving other concerns.

A Theoretical Refinement: The Dual Process Model

After the Rosenblatt et al. (1989) paper was published, a German psychologist, Randolph Ochsmann, reported having difficulty replicating MS effects. In contrast to two brief items, Ochsmann was reminding people of death by leading them through an extensive guided imagery exercise of imagining their own death and interment. This discrepancy led to an important insight into the processes instigated by MS. Realizing that the initial terror management studies invariably included some kind of intervening experience between the MS manipulation and assessment of the dependent measures (e.g., mood scales, experimenter instructions) we posited that terror management defenses do not occur when death is prominent in consciousness but rather when death is highly accessible but no longer in focal attention.

Studies (see Pyszczynski et al., 1999 for a review) initially supported this idea by showing, for example, that MS only leads to world-view defense if there is a non-death-

related task between the MS induction and the dependent measure. They also showed that immediately after an MS induction, the accessibility of death-related thought is low. After a delay, in contrast, death-related thought is no longer in focal attention but becomes high in accessibility. Given people's reluctance to consciously dwell on their own mortality, we hypothesized that, confronted with such thoughts, people initially suppress them. A series of studies supported this reasoning, finding for example that MS led to immediate increases in death-thought accessibility if participants were placed under high cognitive load. Another series of studies further clarified that role of consciousness in MS-induced worldview defenses by showing that worldview defenses emerge in response to unconscious thoughts of death (e.g., when people are exposed to subliminal reminders of mortality).

A dual process model of terror management emerged from this line of research. Explicit thoughts of death instigate direct proximal defenses to remove death-related thoughts from current focal attention. Such "pseudo-rational" mechanisms make death seem like a far-off problem, thereby allowing the individual to stop thinking about it. However, after thoughts of death have been removed from focal attention, death-thought accessibility increases, heightening the potential to experience death-related anxiety. This in turn instigates symbolic distal defenses such as bolstering one's worldview or self-worth. These terror management defenses then bring death thought accessibility down to baseline levels.

Summary

TMT was originally developed to explain two facts of human experience, namely that people have difficulty getting along with those who are different and that people have a trenchant need to feel good about themselves. Research shows that both these propensities are strands in the fabric by which people insulate themselves from deeply rooted fears of mortality. Further, studies have supported a dual process model of the defenses instigated by reminders of mortality. Based on this foundation, the theory has guided investigation of a variety of aspects of human experience, many more so than were envisioned on that snowy day in Utah.

The Grave Matters: Applying TMT to Contemporary Human Concerns

TMT is relevant to a wide range of domains of human behavior. Although we have been accused on occasion of trying to explain [p. 406 ↓] everything with TMT, our position is not that everything stems from the need for terror management, but that most, if not all, significant domains of human behavior are influenced by terror management concerns. Indeed, the theory has provided insights into a wide range of phenomena, including human ambivalence about sex (e.g., Goldenberg et al., 2000), reactions to the handicapped (e.g., Hirschberger, 2006), the operation of stereotypes (e.g., Schimel et al., 1999), academic performance (Landau et al., 2009), altruism (e.g., Jonas et al., 2002), parenthood (e.g., Wisman and Goldenberg, 2005), attachment and close relationships (e.g., Mikulincer et al., 2003), stigmatization (e.g., Salzman, 2001), reactions to nature (e.g., Koole and van den Berg, 2005), attitudes toward women (e.g., Landau et al., 2006a), religion (Greenberg et al., in press), art (e.g., Landau et al., 2006b), film (Sullivan et al., 2010), and even human-android relations (see e.g., MacDorman, 2005).

The theory has also generated some useful insights, following Kurt Lewin, into affairs of particular practical significance, including legal affairs (Arndt et al., 2005a), consumer behavior (Arndt et al., 2004), and mental health (Arndt et al., 2005b). Given space limitations, we will focus on two heavily researched applications of TMT. The first concerns the decisions that people make with regard to the health, and the second concerns the political sphere and our seemingly escalating proclivities for intergroup violence. We will then conclude by considering factors that mitigate the more harmful defensive responses provoked by concerns about mortality.

Understanding the Role of Awareness of Death in Everyday Health Decisions

It seems rather obvious that death does not bode well for an individual's health. Despite this fact, theory and research had largely ignored the deeper psychological significance of concerns about mortality for understanding the decisions people make regarding their physical health. This is surprising given that many health campaigns remind individuals that noncompliance with health recommendations can hasten their demise. What are the possible consequences of such reminders? Goldenberg and Arndt (2008) recently extended TMT with a terror management health model (TMHM) to address such questions.

The TMHM starts with the idea that health-relevant scenarios have varying potential to activate death-related cognition, and then builds from the proximal and distal terror management defenses previously described. When a health situation leads people to explicitly think about death, or thoughts of death are otherwise in focal attention, health decisions are influenced by a proximal motivation to jettison the threatening cognitions from focal attention. For example, an individual may take proactive (e.g., exercise more) or avoidant (e.g., deny perceived risk) steps to render death less of an immediate concern. When thoughts of death are activated implicitly (i.e., outside of focal attention), health-relevant decisions are driven more by a desire to affirm the symbolic value of the self; for example, by pursuing the standards upon which an individual's self-esteem is based, investing in worldview beliefs, or distancing oneself from the creatureliness of the body. Responses to both conscious and unconscious death-related thought can be beneficial or harmful to health, but responses to unconscious death thoughts will be moderated by the individual's worldview and bases of self-worth, whereas responses to conscious death thoughts will be moderated by the perceived efficacy of efforts directed toward reducing the threat of death (see Goldenberg and Arndt, 2008, for a review of the supportive research).

In one study illustrating this basic distinction, participants were reminded of their mortality (or not) and then immediately thereafter or after a delay answered questions [p. 407 ↓] about their intentions to tan. Immediately after thinking about death,

participants reported lower intentions to suntan, suggesting efforts to reduce vulnerability and protect health. However, after a delay (when thoughts of death were accessible but not conscious) participants increased their intentions to suntan, suggesting an effort to bolster esteem by enhancing attractiveness as defined in contemporary society. Thus, conscious thoughts of death led to a health over beauty decision, but unconscious thoughts of death led to a beauty over health decision.

In drawing attention to these distinctions, TMHM helps explain a number of consequences of fear appeals, such as a so-called “boomerang” effect whereby fear appeals (which often highlight the threat of death) can have consequences antithetical to those intended. Whereas the initial response to a fear appeal may reflect efforts to protect one's physical self (which then remove death from consciousness), the delayed effect of unconscious thoughts of death may instigate self-esteem striving that entails more risky behavior.

The TMHM focuses on three general propositions. The first is that people's health decisions will reflect efforts to remove conscious thoughts of death from focal attention. Accordingly, research has found that when confronted with explicit thoughts of death, people react by trying to avoid the threat or by trying to reduce its likelihood with health promoting behaviors (such as increasing exercise). Of course, the critical question is when and for whom a threat avoidance or health promoting response is most likely to emerge. When an individual perceives a health response as effective, maintains optimism, approaches health situations with active coping strategies, or construes vulnerability as low, they respond to conscious thoughts of death with health-oriented rather than threat-avoidance strategies, particularly in death-related health domains (e.g., screenings for breast and skin cancer).

The second direction of research has explored how the unconscious resonance of death-related cognition promotes maintaining, not one's health, but a sense of meaning and self-esteem. For example, thoughts of death outside of focal attention increase both health-defeating (e.g., intentions to suntan) and health-facilitating (e.g., exercising) outcomes depending on situational and dispositional self-esteem contingencies (see Goldenberg and Arndt, 2008). As this suggests, MS can lead to seemingly counterintuitive risky effects on health-relevant outcomes, both for self and others. In one study, after being reminded of mortality, Christian medical students were more

likely to adopt a cautious triage approach to Christian patients complaining of chest pain but adopted a more cavalier approach to Muslim patients reporting identical symptoms (Arndt et al., 2009). This suggests that even people who are more routinely exposed to issues concerning death may still be vulnerable to terror management effects. But why would individuals make such decisions? Such responses help to bolster the symbolic buffers that offer protection from deeply rooted existential fear. Consistent with this analysis, Vess et al. (2009) found that not only did religious fundamentalists respond to reminders of death by endorsing faith-based over medical-based treatments for illness (both for themselves and others), but doing so helped to satiate their need for meaning in life.

This does not imply, however, that unconscious thoughts of death necessarily increase health risk behavior. If individuals derive self-esteem from behaviors conducive to health, they display health-facilitating outcomes across different health domains when death thought accessibility is high. Moreover, in a number of areas (e.g., exercise, smoking cessation, tanning, breast cancer screening), when self-esteem contingencies are experimentally altered in health-conducive ways, people respond to increased death thought accessibility with more protective intentions and behavior. For example, shifting normative appearance standards toward pale skin led actual beachgoers, when distracted from [p. 408 ↓] activated thoughts of death, to request higher SPF sun lotions when offered compensatory samples, especially among those who derive self-esteem from extrinsically defined standards (e.g., Arndt et al., 2009). Indeed, following MS, showing those who smoke to fit in with others a commercial about how smoking reduces one's popularity leads to increased smoking cessation intentions.

The third direction of TMHM research builds from Goldenberg and colleagues' (2000) work on how reminders of our physical, creaturely nature intensify terror management concerns. Confrontations with the physical body threaten the illusion of symbolic significance, and thus our psychological security. Research has shown both reminders of death and of the creaturely aspects of the body influence health-related attitudes and behavior. For example, they increase negative reactions to breastfeeding and pregnant women, increase perceptions of discomfort regarding mammograms, and reduce the thoroughness of breast self-exams. In sum, TMHM research has shown that the awareness of mortality plays a significant role in attitudes and behavior pertinent to cancer detection and prevention as well as to more general healthful living.

Tmt, Politics, and Intergroup Conflict

The awareness of mortality also contributes to intergroup strife, in two primary ways. First, as Becker (1971) noted, the mere presence of groups with worldviews very different from one's own threatens individuals' terror-assuaging faith in their worldview. We have reviewed research showing that reminders of mortality lead to derogation and even aggression against those who criticize the participant's worldview or simply subscribe to a different one.

However, in a chapter called "the Nexus of Unfreedom" from *The Denial of Death*, Becker (1973) elucidated an even more destructive consequence of the fear of death. He argued that the most blood has been spilled not by those with evil motives but by those who are serving the great leader, God, and country. Our mortality requires us to find something bigger than ourselves for our salvation. When that individual or institution designates some other group as evil or that group is perceived as a threat to the terror-assuaging leader, deity, or entity, violent attempts to eradicate that evil threat commonly result.

In his final book, Becker (1975) also noted that no matter what our worldview is, residual death anxiety will exist and people will seek controllable sources of that anxiety to mask its true cause – scapegoats. So the greatest sense of death transcendence is provided by participating in a heroic triumph over evil. In other words, concerns about mortality will draw people toward world-views and leaders who provide them most compellingly with a sense that they are part of something great and have a mission to heroically triumph over those who are evil; rigid worldviews and charismatic leaders that clearly delineate who is good and who is evil will do this best. Becker used this analysis to explain something that otherwise seems inexplicable: the rise of Nazism in Germany culminating in the Holocaust (as well as other similar historical phenomena). Indeed, much of the course of human history seems to have been guided by efforts to heroically triumph over evil in the service of the death-transcending ideologies and leaders in which people have invested.

TMT research has supported these ideas in a number of ways. Reminders of mortality draw people toward charismatic leaders and ideologies that sell a simple vision of the

greatness of the ingroup and the need to rid the world of evildoers. Using descriptions of hypothetical gubernatorial candidates, Cohen et al. (2004) found that MS increased the appeal of a candidate with a charismatic style; that is, one who exuded self-confidence and emphasized the greatness of the state and nation. As the 2004 American presidential election approached, Americans were faced [p. 409 ↓] with a choice that would allow a unique test of the TMT analysis of attraction to leaders. One was George W. Bush, a folksy, at-ease incumbent president who emphasized the greatness of America and the need to rid the world of evildoers, such as the "Axis of Evil." The other was John Kerry, a Democratic challenger with a stiff speaking style who had a complex, sometimes hard to decipher view of issues and who was portrayed by the Republican spin machine as a waffling flip-flopper. Landau et al. (2004b) posited that MS, as well as reminders of terrorism by virtue of activating fears of death, would increase the appeal of Bush and reduce the appeal over Kerry. And in a series of studies prior to the 2004 election, that's exactly what they found. Subsequently, with the help of a Bin Laden tape released days before the election, Bush won re-election even though he had invaded a country and embroiled the US in a protracted war based on false claims.

The Cohen study suggests that the Bush effect resulted from Bush's charismatic style and simple good-versus-evil worldview. However, Bush could have also had an advantage for terror management purposes as the current leader of the nation. In addition, Jost and colleagues (2003) have suggested that conservative rightwing ideologies may serve terror management better than liberal or leftwing ones. However, in the Landau et al. (2004b) studies, despite increasing the preference for Bush, MS did not increase participants' self-reported political conservatism. In addition, a recent study by Kosloff et al. (2010) building on Cohen et al. found that MS leads people to prefer a charismatic hypothetical leader only if that leader also espouses policies that match the individual's pre-existing political orientation, whether conservative or liberal.

So MS seems to move people toward those who espouse more straightforward good-versus-evil ideologies as long as those ideologies fit with the individual's pre-existing worldview. Based on these findings, Pyszczynski et al. (2006) wondered if MS would also increase the appeal of engaging in violence against those designated by one's culture as evil. In their first study, they exposed Iranian college students to MS or a control topic, and then asked them to react to a fellow student who espoused either

suicide bombing against Americans or peaceful approaches to addressing issues with the US. In the control condition, the Iranian students preferred the peaceful fellow student. However, after MS, they preferred the supporter of suicide bombing, and expressed increased interest in joining the cause themselves.

Lest we're tempted to confine such reactions to Iranians, Routledge and Arndt (2008) similarly found that British students after MS reported an increased willingness to die for mother England. Further, people don't just show a proclivity to blow themselves up after MS, but even more so to simply blow up the enemy. Indeed, in a second study, Pyszczynski et al. (2006) manipulated MS and asked conservative and liberal Americans how much they supported extreme military actions, including the use of nuclear weapons, against perceived threats to the US in the Middle East and elsewhere. Compared with the control condition, MS increased conservative participants' advocacy of such violent measures. Similarly, MS increased violent reactions among Israelis upset with the then upcoming 2005 pullout from the Gaza Strip and the northern West Bank (Hirschberger and Ein-Dor, 2006). Finally, supporting the terror management value of death to the designated evil other, Hayes et al. (2008) showed that after threatening Christians' worldviews, informing them that 117 Muslims died in a plane crash actually reduced the accessibility of death-related thought and worldview defense.

Taken together, the TMT literature suggests that the threat of mortality increases negativity toward different others and fuels support for simple "we are good, they are evil" worldviews and desires to rid the world of those designated as evil. Thus, research supports the role of terror management needs in intergroup conflict around the world.

[p. 410 ↓]

Where Do We Go from Here? Insights into Better Ways to Manage Terror

Given the often destructive means that people use to manage the fear of mortality, how can we reduce such effects? Fortunately, a number of lines of research have

addressed this question and suggest ways such responses can be reduced or even reversed. According to Becker, the key is to identify secure, enduring worldviews that provide widely accessible avenues for self-worth to their constituents, but do so in a way that minimizes the costs both to those inside and outside the culture. These insights inspired two broad approaches to mitigating harmful terror management that have been examined empirically.

Fortifying the Shield: Psychological Buffers against Thoughts of Mortality

The earliest statements of TMT posited that, because self-esteem serves to buffer anxiety, boosting self-esteem or having high dispositional self-esteem should render people less vulnerability to the consequences of awareness of mortality. Harmon-Jones et al. (1997) and others have supported the specific hypothesis that bolstering an individual's self-worth would enable the person to contemplate mortality without experiencing subsequent increases in death thought accessibility and associated defensive responses. As long as the threat that other people pose does not undermine their own feelings of self-worth (Arndt and Greenberg, 1999), self-esteem is capable of offering people psychological protection against responding defensively to reminders of death.

People can also derive existential protection through their attachments with close others. Rank (1941/1958) originally proposed that interpersonal relationships have become an especially important security blanket as dominant religious doctrines have lost some of their luster. In accord with this insight, Mikulincer and others have shown that people who have secure attachments to close others when reminded of death are less likely to respond defensively and advocate violence (e.g., Mikulincer et al., 2003; Weise et al., 2008).

While both relational attachment and self-esteem (and other buffers we don't have the space to touch on here) reduce consequences of MS, these critical bases of psychological security are often hard to sustain over the lifespan. When they falter, problems with anxiety, depression, and substance abuse are likely. Thus, the more

durable the bases of self-worth and attachments, the less destructive an individual's efforts at terror management are likely to be.

Worldviews Espousing Tolerance, Compassion, and Open-Mindedness

Another tack for mitigating harmful forms of terror management focuses not on fortifying the existential shield, but on funneling the routes that responses take in more constructive directions. This can be achieved by appealing to worldviews that espouse tolerance and open-mindedness. If an individual values tolerance and MS motivates effort to live up to important values, then MS should increase tolerance. Accordingly, Greenberg and colleagues (1992a) showed that when tolerance was primed or dispositionally important, participants did not respond to MS with increased derogation of those who threatened their beliefs. Such findings of course are hopeful, as people can socialize their children into valuing acceptance of others.

Many components of culture contain such prescriptions, though the message often gets lost. For instance, could religiously advocated compassion encourage more peaceful coexistence with those who are different? To find out, Rothschild et al. (2009) investigated how priming compassionate religious values might redirect the effects of MS on support for violent worldview defense among religious fundamentalists. In these studies, using [p. 411 ↓] both American samples as well as a Shia population in Iran, participants were reminded of death or a control topic, and then were presented with either compassionate or neutral religious scripture, or compassionate or neutral secular adages, followed by a measure of support for extreme military responses (against those who challenge the US or toward the US in the case of the Shia). Religious fundamentalists in the control condition were supportive of violent worldview defense regardless of type of teaching. However, fundamentalists reminded of death and exposed to compassionate religious teachings demonstrated a significant decrease in their support for violence. Presumably the death reminder motivated an increased need to live up to compassionate values when they were linked to the worldview.

Can tolerance and compassion be accentuated without explicit priming? One possibility involves people reconceptualizing ingroup identifications to recognize our common humanity with people the world over. For example, Motyl and colleagues (submitted) reminded American participants of their death or a control topic, and then exposed them to pictures of common non-American families from diverse regions of the world, stereotypically American families, or a group of unrelated people. In the neutral and American families conditions, MS elicited an increase in anti-Arab prejudice; however, when reminded of the common humanity shared by all people, MS significantly reduced anti-Arab prejudice. In a second study, Americans were reminded either of death or a control topic, and then presented with brief stories of common childhood experiences ostensibly written by either Americans or people from diverse areas of the world. When presented with American childhood experiences, MS increased hostile prejudice against immigrants; however, when those same childhood experiences were attributed to a foreign author, this negative effect was eliminated.

Encouraging people to consider our shared humanity is a promising avenue to reducing destructive effects, particularly given the increasing globalization of the world. Such goals may also be facilitated by more generally stimulating open-minded and flexible thinking. By virtue of its utilization of open-minded thinking, creative action may be one way to encourage more acceptance of different others.

The initial research examining creativity and terror management (see Greenberg et al., 2008, for a review) built from the theorizing of Otto Rank (e.g., 1932). Creative engagement, by separating the individual from conventional thinking, threatens the person's valued place within the security-providing worldview. Therefore, when reminded of mortality, creative action can lead to feelings of guilt (an emotion that reflects a desire for social reparation). But fortified with an enhanced sense of such social connection, people can engage in creativity after being reminded of death without experiencing guilt, and can then reap the more positive psychological effects of such activity. Accordingly, after being reminded of death, people are more creative if the product is directed toward communal benefit, but less creative if the product is directed toward individual gain.

This work also suggests that as people are faced with managing existential fears, creativity has the potential to facilitate more optimal engagement with life (Routledge

and Arndt, 2009). One study found that socially validated creativity inspires a more open-minded orientation which reduces people's tendency to manage terror by derogating those with conflicting beliefs. Another set of studies showed that priming the cultural value of creativity after MS can even increase willingness to expose oneself to ideas that run counter to prevailing cultural beliefs. This and other research suggests that people can manage the awareness of death in ways that actually embrace the rich diversity of perspectives that the world offers, the possibility of authentic self-transcendence, and intrinsic goal pursuit (see, for example, Lykins et al., 2007).

Finally, individual differences in need for structure may play a role in how people cope [p. 412 ↓] with their mortality concerns. Vess, Routledge et al. (2009) showed in a series of studies that while those high in need for structure prefer more rigid routes to epistemic knowledge (as in the work of Landau et al., 2004a, 2006a), those low in need for structure may utilize more explorative forms of discovery and integrative processing to extract meaning and significance in their life. Thus, again, much of the potential for reducing destructive forms of terror management seems to go back to the content of the worldview to which an individual subscribes. When one's worldview prescribes prosocial behavior, flexible thinking, or tolerance and compassion, constructive responses to the human existential predicament are likely.

Long Day's Journey into Night

TMT has brought the problem of death into social psychology, and we hope that you agree that has been a valuable (albeit not pleasant to think about) contribution. The threat of death lurks behind most, if not all, the things people care about: health, economic wellbeing, the environment, terrorism, war, close relationships, aging, and spirituality. TMT work has helped clarify how the awareness of mortality contributes to attitudes and behavior regarding these topics and many more as well. In our defiance of our mortality, we all urgently defend our cherished beliefs and strive hard to make the greatest mark on the world we can. TMT research has shown that these urgings contribute to the most noble and ignoble forms of human behavior.

TMT work has also raised many questions currently being considered. How do defenses aroused by conscious and unconscious concerns relate to each other?

What does TMT work imply about the basic nature of the unconscious? What brain regions contribute to the awareness of mortality and the defenses it instigates? What are the effects of sustained heightened mortality salience over time; for example, for professionals such as oncologists and morticians? One study showed that workers who perform death ceremonies in India had stronger pro-India bias than other workers, independent of an MS induction, whereas the other workers were as strongly pro-India as the death ceremony workers only if they were first induced to consider their mortality (Fernandez et al., 2010). But much more research on this issue is needed. Are there particular components of cultural worldviews that are particularly important for terror management? Are there certain beliefs and values that are common to all worldviews because of the way evolutionary has shaped our basic motivations, ways of thinking, and emotions (e.g., empathy, desire for justice)? Are there other motivations besides terror management which contribute to the desires to preserve faith in a meaningful world and one's own significance?

We also like to think that TMT has contributed to a general broadening of the field, a greater awareness of the important roles of culture, of unconscious motivations and processes, of emotions, and of core human concerns in social behavior. One aspect of this has been the emergence of the subfield within social psychology known as experimental existential psychology (XXP; Greenberg et al., 2004; Pyszczynski et al., 2010). This subfield focuses on the impact on thought, feeling, and behavior of five major existential concerns: death, meaning, identity, isolation, and freedom. We are hopeful that XXP will continue to progress toward a comprehensive understanding of these core facets of the human condition and their roles in the social phenomena that matter to us most; such a prospect gives us a sense of fundamental satisfaction.

References

Ahrendorf, P.J. (2000) The fear of death and the longing for immortality: Hobbes and Thucydides on human nature and the problem of anarchy . The American Political Science Review , 94, 579–593. <http://dx.doi.org/10.2307/2585832>

Arndt, J., Cox, C.R., Goldenberg, J.L., Vess, M., Routledge, C. and Cohen, F. (2009) Blowing in the (social) wind: Implications of extrinsic esteem contingencies for terror

management and health . *Journal of Personality and Social Psychology* , 96, 1191–1205. <http://dx.doi.org/10.1037/a0015182>

Arndt, J. and Greenberg, J. (1999) The effects of a self-esteem boost and mortality salience on responses to boost relevant and irrelevant worldview threats . *Personality and Social Psychological Bulletin* , 25, 1331–1341. <http://dx.doi.org/10.1177/0146167299259001>

Arndt, J., Greenberg, J., Solomon, S., Pyszczynski, T. and Simon, L. (1997) Suppression, accessibility of death-related thoughts, and cultural worldview defense: Exploring the psychodynamics of terror management . *Journal of Personality and Social Psychology* , 73, 5–18. <http://dx.doi.org/10.1037/0022-3514.73.1.5>

Arndt, J., Lieberman, J.D., Cook, A. and Solomon, S. (2005a) Terror management in the courtroom: Exploring the effects of mortality salience on legal decision-making . *Psychology, Public Policy, and Law* , 11, 407–438. <http://dx.doi.org/10.1037/1076-8971.11.3.407>

Arndt, J., Routledge, C., Cox, C.R. and Goldenberg, J.L. (2005b) The worm at the core: A terror management perspective on the roots of psychological dysfunction . *Applied and Preventative Psychology* , 11, 191–213. <http://dx.doi.org/10.1016/j.appsy.2005.07.002>

Arndt, J., Solomon, S., Kasser, T. and Sheldon, K.M. (2004) The urge to splurge: A terror management account of materialism and consumer behavior . *Journal of Consumer Psychology* , 14, 198–212. http://dx.doi.org/10.1207/s15327663jcp1403_2

Arndt, J., Vess, M., Cox, C.R., Goldenberg, J.L. and Lagle, S. (2009) The psychosocial effect of personal thoughts of mortality on cardiac risk assessment . *Medical Decision Making* , 29, 175–181. <http://dx.doi.org/10.1177/0272989X08323300>

Baldwin, J. (1963) *The Fire Next Time* . New York: Random House.

Becker, E. (1971) *The Birth and Death of Meaning* , 2nd Edition. New York: Free Press.

Becker, E. (1973) *The Denial of Death* . New York: Academic Press.

- Becker, E. (1975) Escape from Evil . New York: Academic Press.
- Cohen, F., Solomon, S., Maxfield, M., Pyszczynski, T. and Greenberg, J. (2004) Fatal attraction: The effects of mortality salience on evaluations of charismatic, task-oriented, and relationship-oriented leaders . Psychological Science , 15, 846–851. <http://dx.doi.org/10.1111/j.0956-7976.2004.00765.x>
- Fernandez, S., Castano, E. and Singh, I. (2010) Managing death in the burning grounds of Varanasi, India: A terror management investigation . Journal of Cross-Cultural Psychology , 41, 182–194. <http://dx.doi.org/10.1177/0022022109354376>
- Friedman, M. and Rholes, W.S. (2007) Successfully challenging fundamentalist beliefs results in increased death awareness . Journal of Experimental Social Psychology , 43, 794–801. <http://dx.doi.org/10.1016/j.jesp.2006.07.008>
- Goldenberg, J.L. and Arndt, J. (2008) The implications of death for health: A terror management health model for behavioral health promotion . Psychological Review , 115, 1032–1053. <http://dx.doi.org/10.1037/a0013326>
- Goldenberg, J.L., McCoy, S.K., Pyszczynksi, T., Greenberg, J. and Solomon, S. (2000) The body as a source of self-esteem: The effects of mortality salience on identification with one's body, interest in sex, and appearance monitoring . Journal of Personality and Social Psychology , 79, 118–130. <http://dx.doi.org/10.1037/0022-3514.79.1.118>
- Greenberg, J., ed. , Koole, S., ed. and Pyszczynski, T. (eds) (2004) Handbook of Experimental Existential Psychology . New York: Guilford Press.
- Greenberg, J., Landau, M.J., Solomon, S. and Pyszczynski, T. (in press) The case for terror management as the primary psychological function of religion. In D. Wulff (ed.) Handbook of the Psychology of Religion . London: Oxford University Press.
- Greenberg, J., Pyszczynski, T. and Solomon, S. (1986) The causes and consequences of a need for self-esteem: A terror management theory . In R.F. Baumeister (ed.), Public Self and Private Self , pp. 189–212. New York: Springer-Verlag. http://dx.doi.org/10.1007/978-1-4613-9564-5_10

Greenberg, J., Simon, L., Pyszczynski, T., Solomon, S. and Chatel, D. (1992a) Terror management and tolerance: Does mortality salience always intensify negative reactions to others who threaten one's worldview? *Journal of Personality and Social Psychology* , 63, 212–220. <http://dx.doi.org/10.1037/0022-3514.63.2.212>

Greenberg, J., Solomon, S. and Arndt, J. (2008) A uniquely human motivation: Terror management . In J. Shah, ed. and W. Gardner (eds), *Handbook of Motivation Science* , pp. 113–134. New York: Guilford.

Greenberg, J., Solomon, S., Pyszczynski, T., Rosenblatt, A., Burling, J., Lyon, D., Pinel, E. and Simon, L. (1992b) Assessing the terror management analysis of self-esteem: Converging evidence of an anxiety-buffering function . *Journal of Personality and Social Psychology* , 63, 913–922. <http://dx.doi.org/10.1037/0022-3514.63.6.913>

Harmon-Jones, E., Simon, L., Greenberg, J., Pyszczynski, T., Solomon, S. and McGregor, H. (1997) Terror management theory and self-esteem: Evidence that increased self-esteem reduces mortality salience effects . *Journal of Personality and Social Psychology* , 72, 24–36. <http://dx.doi.org/10.1037/0022-3514.72.1.24>

Hayes, J., Schimel, J. and Williams, T.J. (2008) Fighting death with death: The buffering effects of learning that worldview violators have died after worldview threat . *Psychological Science* , 19, 501–507. <http://dx.doi.org/10.1111/j.1467-9280.2008.02115.x>

HirschbergerG. (2006) Terror management and attributions of blame to innocent victims: Reconciling compassionate and defensive responses . *Journal of Personality and Social Psychology* , 91, 832–844. <http://dx.doi.org/10.1037/0022-3514.91.5.832>

Hirschberger, G. and Ein-Dor, T. (2006) Defenders of a lost cause: Terror management and violent resistance to the disengagement plan . *Personality and Social Psychology Bulletin* , 32, 761–769. <http://dx.doi.org/10.1177/0146167206286628>

Jonas, E., Schimel, J., Greenberg, J. and Pyszczynski, T. (2002) The Scrooge effect: Evidence that mortality salience increases prosocial attitudes and behavior . *Personality and Social Psychology Bulletin* , 28, 1342–1353. <http://dx.doi.org/10.1177/014616702236834>

Jost, J.T., Glaser, J., Kruglanski, A.W. and Sulloway, F. (2003) Political conservatism as motivated social cognition . *Psychological Bulletin* , 129, 339–375. <http://dx.doi.org/10.1037/0033-2909.129.3.339>

Koole, S.L. and van den Berg, A.E. (2005). Lost in the wilderness: Terror management, action orientation, and nature evaluation . *Journal of Personality and Social Psychology* , 88, 1014–1028. <http://dx.doi.org/10.1037/0022-3514.88.6.1014>

Kosloff, S., Greenberg, J., Weise, D. and Solomon, S. (2010) The effects of reminders of death on political preferences: The roles of charisma and political orientation . *Journal of Experimental Social Psychology* , 46, 139–145. <http://dx.doi.org/10.1016/j.jesp.2009.09.002>

Landau, M.J., Goldenberg, J., Greenberg, J., Gillath, O., Solomon, S., Cox, C., Martens, A. and Pyszczynski, T. (2006a) The siren's call: Terror management and the threat of men's sexual attraction to women . *Journal of Personality and Social Psychology* , 90, 129–146. <http://dx.doi.org/10.1037/0022-3514.90.1.129>

Landau, M.J., Greenberg, J. and Rothschild, Z. (2009) Motivated cultural worldview adherence and culturally loaded test performance . *Personality and Social Psychology Bulletin* , 35, 442–453. <http://dx.doi.org/10.1177/0146167208329630>

Landau, M.J., Greenberg, J., Solomon, S., Pyszczynski, T. and Martens, A. (2006b) Windows into nothingness: Terror management, meaninglessness, and negative reactions to modern art . *Journal of Personality and Social Psychology* , 90, 879–892. <http://dx.doi.org/10.1037/0022-3514.90.6.879>

Landau, M.J., Johns, M., Greenberg, J., Pyszczynski, T., Solomon, S. and Martens, A. (2004a) A function of form: Terror management and structuring of the social world . *Journal of Personality and Social Psychology* , 87, 190–210. <http://dx.doi.org/10.1037/0022-3514.87.2.190>

Landau, M.J., Solomon, S., Greenberg, J., Cohen, F., Pyszczynski, T., Arndt, J., Miller, C.H., Ogilvie, D.M. and Cook, A. (2004b) Deliver us from evil: The effects of mortality salience and reminders of 9/11 on support for President George

W. Bush . *Personality and Social Psychology Bulletin* , 30, 1136–1150. <http://dx.doi.org/10.1177/0146167204267988>

Lykins, E.L.B., Segerstrom, S.C., Averill, A.J., Evans, D.R. and Kemeny, M.E. (2007) Goal shifts following reminders of mortality: Reconciling posttraumatic growth and terror management theory . *Personality and Social Psychology Bulletin* , 33, 1088–1099. <http://dx.doi.org/10.1177/0146167207303015>

MacDorman, K.F. (2005) Mortality salience and the uncanny valley . Proceedings of the 5th IEEE-RAS International Conference on Humanoid Robots , Tsukuba, Japan, 399–405. Doi: 10.1109/ICHR.2005.1573600. <http://dx.doi.org/10.1109/ICHR.2005.1573600>

McGregor, I., Zanna, M.P., Holmes, J.G. and Spencer, S.J. (2001) Compensatory conviction in the face of personal uncertainty: Going to extremes and being oneself . *Journal of Personality and Social Psychology* , 80, 472–488. <http://dx.doi.org/10.1037/0022-3514.80.3.472>

Mikulincer, M., Florian, V. and Hirschberger, G. (2003) The existential function of close relationships: Introducing death into the science of love . *Personality and Social Psychology Review* , 7, 20–40. http://dx.doi.org/10.1207/S15327957PSPR0701_2

Motyl, M., Pyszczynski, T., Cox, C., Siedel, A., Maxfield, M. and Weise, D.R. (submitted) We are family: The effects of mortality salience and priming a sense of common humanity on intergroup prejudice.

Pyszczynski, T., Abdollahi, A., Solomon, S., Greenberg, J., Cohen, F. and Weise, D. (2006). Mortality salience, martyrdom, and military might: The Great Satan versus the Axis of Evil . *Personality and Social Psychology Bulletin* , 32, 525–537. <http://dx.doi.org/10.1177/0146167205282157>

Pyszczynski, T., Greenberg, J., Koole, S. and Solomon, S. (2010) Experimental existential psychology: Coping with the facts of life . In S. Fiske, ed. , D. Gilbert, ed. , and G. Lindzey (eds), *Handbook of Social Psychology*, Vol. 1 , 5th Edition, pp. 724–757. London: John Wiley and Sons.

Pyszczynski, T., Greenberg, J. and Solomon, S. (1999) A dual-process model of defense against conscious and unconscious death-related thoughts: An extension of terror management theory . *Psychological Review* , 106, 835–845. <http://dx.doi.org/10.1037/0033-295X.106.4.835>

Rank, O. (1932/1989) Art and Artist: Creative Urge and Personality Development . New York: Knopf.

Rank, O. (1941/1958) Beyond Psychology . New York: Dover.

Rosenblatt, A., Greenberg, J., Solomon, S., Pyszczynski, T. and Lyon, D. (1989) Evidence for terror management theory I: The effects of mortality salience on reactions to those who violate or uphold cultural values . *Journal of Personality and Social Psychology* , 57, 681–690. <http://dx.doi.org/10.1037/0022-3514.57.4.681>

Rothschild, Z.K., Abdollahi, A. and Pyszczynski, T. (2009) Does peace have a prayer? The effect of mortality salience, compassionate values, and religious fundamentalism on hostility toward outgroups . *Journal of Experimental Social Psychology* , 45, 816–827. <http://dx.doi.org/10.1016/j.jesp.2009.05.016>

Routledge, C. and Arndt, J. (2008) Self-sacrifice as self-defense: Mortality salience increases efforts to affirm a symbolic immortal self at the expense of the physical self . *European Journal of Social Psychology* , 38, 531–541. <http://dx.doi.org/10.1002/ejsp.442>

Routledge, C. and Arndt, J. (2009) Creative terror management: Creativity as a facilitator of cultural exploration after mortality salience . *Personality and Social Psychology Bulletin* , 35, 493–505. <http://dx.doi.org/10.1177/0146167208329629>

Salzman, M.B. (2001) Cultural trauma and recovery: Perspectives from terror management theory . *Trauma Violence and Abuse* , 2, 172–191. <http://dx.doi.org/10.1177/1524838001002002005>

Schimel, J., Hayes, J., Williams, T.J. and Jahrig, J. (2007) Is death really the worm at the core? Converging evidence that worldview threat increases death-thought

accessibility . *Journal of Personality and Social Psychology* , 92, 789–803. <http://dx.doi.org/10.1037/0022-3514.92.5.789>

Schimel, J., Simon, L., Greenberg, J., Pyszczynski, T., Solomon, S., Waxmonski, J. and Arndt, J. (1999) Support for a functional perspective on stereotypes: Evidence that mortality salience enhances stereotypic thinking and preferences . *Journal of Personality and Social Psychology* , 77, 905–926. <http://dx.doi.org/10.1037/0022-3514.77.5.905>

Solomon, S., Greenberg, J. and Pyszczynski, T. (1991) A terror management theory of social behavior: The psychological functions of self-esteem and cultural worldviews . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 24, 93–159. New York: Academic Press.

Sullivan, D., Greenberg, J. and Landau, M.J. (2010) Toward a new understanding of two films from the dark side: Terror management theory applied to Rosemary's Baby and Straw Dogs . *Journal of Popular Film and Television* , 37, 42–51.

van den Bos, K. (2001) Uncertainty management: The influence of uncertainty salience on reactions to perceived procedural fairness . *Journal of Personality and Social Psychology* , 80, 931–941. <http://dx.doi.org/10.1037/0022-3514.80.6.931>

Vess, M., Arndt, J., Cox, C.R., Routledge, C. and Goldenberg, J.L. (2009) The terror management of medical decisions: The effect of mortality salience and religious fundamentalism on support for faith-based medical intervention . *Journal of Personality and Social Psychology* , 97, 334–350. <http://dx.doi.org/10.1037/a0015545>

Vess, M., Routledge, C., Landau, M. and Arndt, J. (2009) The dynamics of death and meaning: The effects of death-relevant cognitions and personal need for structure on perceptions of meaning in life . *Journal of Personality and Social Psychology* , 97, 728–744. <http://dx.doi.org/10.1037/a0016417>

Weise, D., Pyszczynski, T. Cox, C., Arndt, J., Greenberg, J., Solomon, S. and Kosloff, S. (2008) Interpersonal politics: The role of terror management and attachment processes in political preferences . *Psychological Science* , 19, 448–455. <http://dx.doi.org/10.1111/j.1467-9280.2008.02108.x>

Wisman, A. and Goldenberg, J.L. (2005) From the grave to the cradle: Evidence that mortality salience engenders a desire for offspring . Journal of Personality and Social Psychology , 89, 46–61. <http://dx.doi.org/10.1037/0022-3514.89.1.46>

<http://dx.doi.org/10.4135/9781446249215.n20>

Handbook of Theories of Social Psychology: Volume 1

Self-Determination Theory

Contributors: Edward L. Deci & Richard M. Ryan

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Self-Determination Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n21>

Print pages: 416-438

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n21>

[p. 416 ↓]

Chapter 20: Self-Determination Theory

Edward L. Deci, ed. and Richard M. Ryan, ed.

Abstract

Self-determination theory (SDT) is an empirically derived theory of human motivation and personality in social contexts that differentiates motivation in terms of being autonomous and controlled. Work leading to the theory began with experiments examining the effects of extrinsic rewards on intrinsic motivation. During more than thirty years since the initial studies, we have developed five mini-theories to address different, though related, issues: the effects of social environments on intrinsic motivation; the development of autonomous extrinsic motivation and self-regulation through internalization and integration; individual differences in general motivational orientations; the functioning of fundamental universal psychological needs that are essential for growth, integrity, and wellness; and the effects of different goal contents on well-being and performance. We have subsequently used SDT and its mini-theories to guide and interpret research on many new issues, including motivation and wellness across cultures, close relationships, enhancement and depletion of energy and vitality, and the roles of both mindful awareness and nonconscious processes in behavioral regulation. Although much of SDT was developed through laboratory experiments, it is also supported by a great deal of applied research using both field studies and clinical trials to address significant social issues. We briefly mention some of that work, especially related to health behavior change, education, psychotherapy, work motivation, sport and exercise, and prosocial behaviors.

Introduction

The central focus of social psychology has long been the effects of social environments on people's attitudes, values, motivations, and behaviors, and there can be little doubt that environmental forces have an enormous impact on these outcomes. Many social psychological theories have tended, either implicitly or explicitly, to view learning – that is, acquiring the attitudes, values, motivations, and behaviors – primarily in terms of social environments teaching individuals what to think, value, need, and do. This view has been called the “standard social science model” of a relatively plastic human nature, molded by its social contexts (see Tooby and Cosmides, 1992). In developmental psychology, this [p. 417 ↓] perspective is most clearly conveyed in social learning theory, where modeling and reinforcements are the principal mechanisms of learning and growth. In social psychology, social cognition, and cultural relativism, such models are frequently employed to explain situationally influenced cognitions and behaviors.

The social psychology of self-determination theory (SDT) has also focused on the influences of social environments on attitudes, values, motivations, and behaviors both developmentally and in current situations; however, it has taken a quite different approach to these issues. Specifically, SDT assumes that the human organism is evolved to be inherently active, intrinsically motivated, and oriented toward developing naturally through integrative processes. These qualities need not be learned; they are inherent in human nature. Still, they develop over time, play a central role in learning, and are affected by social environments.

For these natural, active processes of intrinsic motivation and integration to operate effectively toward healthy development and psychological well-being, human beings need particular nutriments – both biological and psychological (Ryan, 1995). In the relative absence of such nutriments, these natural processes will be impaired, resulting in experiences, development, and behaviors that are less than optimal. Within SDT, we focus primarily on psychological nutriments and their dynamics within social environments, although biological supports, as well as inherent individual differences, also play important roles.

According to SDT, based on decades of empirical work, there are at least three universal psychological needs – specifically, needs for competence, autonomy, and relatedness – that are essential for optimal development and functioning. Unlike some evolutionary perspectives, we understand these needs as underlying the adaptive organization of behavior and being supported by many individual adaptations, rather than themselves being functionally specific or modular “add-ons” (Deci and Ryan, 2000). As will be apparent throughout the chapter, the starting point of natural activity, intrinsic motivation, the integrative tendency, and fundamental psychological needs has led to many instances where the predictions of SDT have diverged dramatically from those of other prominent social-psychological theories, and we will point out a few of them in this chapter.

The Development of SDT

SDT evolved out of research on the effects of extrinsic rewards on intrinsic motivation. In the first published studies (Deci, 1971), college students were paid for working on intrinsically interesting puzzles and the monetary rewards undermined their intrinsic motivation for the activity. This was followed by more than 100 similar studies (see Deci et al., 1999) confirming this controversial idea that rewards do not always motivate subsequent persistence; indeed they can undermine intrinsic motivation. We have used the attributional concept *perceived locus of causality* (de Charms, 1968; Heider, 1958) as part of our account of that effect and of other changes in intrinsic motivation, but we have also linked intrinsic motivation, and the social-contextual effects on it, to the basic human needs for competence and self-determination (i.e., autonomy). Intrinsic motivation was considered an inherent characteristic of human beings and was viewed as the prototype of psychological freedom or self-determination. It could be either undermined or enhanced depending on whether the social environment supported or thwarted the needs for competence and self-determination. If a reward or other external event such as threat of punishment (Deci and Cascio, 1972), positive feedback (Deci, 1971), competition (Deci and Betley et al., 1981), or choice (Zuckerman et al., 1978) were expected to thwart these basic needs, it was predicted to prompt an external [p. 418 ↓] perceived locus of causality and undermine intrinsic motivation; but if the event were expected to support these basic needs, it was predicted to prompt

an internal perceived locus of causality and enhance intrinsic motivation. Monetary rewards, threats, and competition were predicted to thwart autonomy, and such events did typically undermine intrinsic motivation (e.g., Deci, Betley et al., 1981; Deci et al., 1999). In contrast, positive feedback and choice were predicted to enhance experiences of competence and self-determination, fostering greater intrinsic motivation, and results have confirmed this as well (Deci et al., 1999; Zuckerman et al., 1978).

Our theoretical postulate linking environmental factors to fundamental human needs as a basis for explaining effects of the social environment on intrinsic motivation stood in sharp contrast to the views of other psychologists examining intrinsic motivation. For example, Lepper et al. (1973) considered intrinsic motivation to be a postbehavior self-attribution, and they used Bern's (1972) self-perception theory to explain the undermining of intrinsic motivation by tangible rewards – that is, people simply attributed less intrinsic motivation to themselves because there was overjustification for doing the activity.

Cognitive Evaluation Theory

As the work progressed, it became increasingly necessary to consider the dynamic interplay of autonomy and competence needs within social contexts in order to explain the increasingly complex experimental phenomena such as the finding that performance-contingent rewards, which are given for doing well on a task (e.g., doing better than the 80th percentile), were less detrimental than task-contingent rewards, which are given for doing or completing a task (Ryan et al., 1983). Accordingly, we (Deci and Ryan, 1980) reviewed the extant literature and introduced a formal mini-theory, labeled *cognitive evaluation theory* (CET), to explain the effects of extrinsic factors on intrinsic motivation.

CET specified two processes through which intrinsic motivation can be affected. First, to the extent that events such as rewards lead to an external perceived locus of causality and thwart the autonomy or self-determination need, the events will undermine intrinsic motivation; whereas, to the extent that events such as choice lead to an internal perceived locus of causality and support the autonomy need, they will enhance intrinsic motivation. The second process specifies that events such as positive feedback

that lead to perceived competence by supporting the competence need will enhance intrinsic motivation, whereas events such as negative feedback that lead to perceived incompetence will undermine intrinsic motivation. However, the positive feedback must be for an autonomously motivated activity (Pritchard et al., 1977) or within an autonomy-supportive context (Ryan, 1982) in order for it to enhance intrinsic motivation.

Finally, CET specified that social-environmental events such as rewards or feedback have two aspects relevant to intrinsic motivation. The first is a *controlling* aspect that pressures people to think, feel, or behave in specific ways, thus prompting an external perceived locus of causality, thwarting autonomy, decreasing intrinsic motivation, and leaving motivation primarily controlled rather than autonomous. The second aspect is the *informational* aspect that conveys competence information within the context of some autonomy support. When this aspect affirms people's competence for an autonomous activity, it supports the competence need and to some extent the autonomy need, thus enhancing intrinsic motivation. However, when it prompts perceived incompetence and thwarts the competence need, it decreases intrinsic motivation. In fact, if the competence information is sufficiently negative, indicating that people are too incompetent to attain desired outcomes, it could undermine both intrinsic and extrinsic motivation, leaving individuals *amotivated* (i.e., without intention or [p. 419 ↓] motivation toward the activity). The effect of an event on intrinsic motivation will depend on the relative salience of the two aspects. The CET proposition about the two aspects of rewards provided an explanation for such phenomena as tangible rewards undermining intrinsic motivation and verbal rewards (i.e., positive feedback) enhancing it. It also explained, for example, why performance-contingent rewards, although they significantly undermined intrinsic motivation, were not as detrimental as task-contingent rewards: because the controlling aspect of both types of rewards is similarly salient, but the informational aspect is more salient in performance-contingent rewards than in task-contingent rewards (Ryan et al., 1983).

Reward Effects Revisited

The research on reward effects was very controversial from its first appearance, and to some extent the controversy continued for decades. People not fond of the finding that tangible rewards tend to undermine intrinsic motivation argued that there were

confounds in the methods (e.g., Calder and Staw, 1975), that there were behavioral explanations of the phenomena that were more valid than CET's cognitive-motivational ones (e.g., Scott, 1975), and that the findings provided no real reason to refrain from using rewards as a primary motivational strategy in education and elsewhere (Eisenberger and Cameron, 1996). These critiques led us to do a meta-analysis of 128 experiments on the effects of extrinsic rewards on intrinsic motivation (Deci et al., 1999). The meta-analysis strongly confirmed what we had long been saying, namely that: (1) positive feedback enhances intrinsic motivation; (2) tangible rewards undermine intrinsic motivation; and (3) both task-contingent rewards and performance-contingent rewards decrease intrinsic motivation, but that unexpected rewards and rewards that do not require doing the target task do not undermine intrinsic motivation for it. Further, CET provided a full account of the complex set of findings that emerged from the meta-analysis.

The Social Ambience

There was another important set of findings that emerged from the CET research, namely that the general interpersonal ambience of a situation such as a classroom or a work group can be characterized as either autonomy supportive or controlling. For example, Deci, Schwartz et al. (1981) found that when elementary-school teachers created an autonomy-supportive classroom ambience their students evidenced greater intrinsic motivation as well as greater competence need satisfaction, than when the teachers created a controlling ambience, because, in the former context, the students felt free to develop their own sense of competence. Further Deci et al. (1989) found that when managers were more autonomy supportive their subordinates were more satisfied and trusting than when the managers were more controlling.

In laboratory experiments, the social ambience was manipulated to examine the effects of various external events when administered in an autonomy-supportive versus controlling context. For example, Ryan (1982) found that when positive feedback was given within an autonomy-supportive ambience it tended to enhance intrinsic motivation, as had been found earlier, but when it was given within a controlling ambience it decreased intrinsic motivation, thus confirming that positive feedback enhances intrinsic motivation only if it is accompanied by some autonomy support. (In short, forced

competence development does not enhance intrinsic motivation.) Similarly, Ryan et al. (1983) found that when performance-contingent monetary rewards were administered controllingly they undermined intrinsic motivation, as had been found earlier, but when they were administered in an autonomy-supportive context they enhanced intrinsic motivation relative to no rewards and no feedback, although these rewards still led to less intrinsic motivation than just positive feedback comparable to what was implicit in the performance-contingent rewards (e.g., “you did better than 80 percent of the other participants”). [p. 420 ↓] Koestner et al. (1984) found further that when limits were set on children in an autonomy-supportive context the limits had a positive effect on subsequent motivation, but when they were set in a controlling context their effect was negative.

Summary

The CET work was able to specify the types of external events that would, on average, enhance intrinsic motivation (e.g., positive feedback and choice) versus diminish it (e.g., tangible rewards and competition). The research was also able to characterize autonomy-supportive versus controlling social contexts and use those to predict the intrinsic motivation of people within them. Finally, it explained how the social context or ambience could interact with external events to moderate their results on intrinsic motivation.

Additional Mini-Theories

As we were writing the review that drew together the CET research (Deci and Ryan, 1980), we began thinking about two new questions. First, would it make sense to have an individual difference concept that paralleled the “state” ideas of autonomous motivation, controlled motivation, and amotivation? This seemed important because surely enduring person factors and not just concurrent social-contextual factors affect people's motivation and behavior at any given time. The second question was whether extrinsically motivated behaviors, which the studies showed typically thwarted the autonomy need and undermined intrinsic motivation, could be performed autonomously,

and, if so, how autonomous extrinsic motivation could be facilitated. This also seemed like a very important question, because daily behaviors involve for most of us a hefty portion of extrinsically motivated activities, some of which are enacted autonomously and others of which are clearly controlled and characterized by alienation.

Causality Orientations Theory

The first of these two new questions led to the concept of *causality orientations* – with the autonomous, controlled, and impersonal causality orientations – as individual differences (Deci and Ryan, 1985a). Everyone is said to have each of the orientations to some degree, so any or all of them can be used to predict outcomes. The autonomy orientation refers both to orienting toward internal and external cues in a way that gives them an autonomy-supportive or informational significance and also to being more autonomous in general across domains and time. The controlled orientation refers to interpreting cues as controls and demands and to being controlled in general at the person level. Finally, the impersonal orientation refers to orienting toward cues as indicators of incompetence and to being generally amotivated.

The causality orientations concept and its psychometric scale have worked effectively in predicting variance in a range of adult outcomes. For example, the autonomy orientation has been positively associated with self-actualization, self-esteem, more choiceful self-disclosure, and supporting autonomy in others; the controlled orientation has been positively associated with public self-consciousness, the type-A coronary prone behavior pattern, inconsistency in attitudes and behaviors, and greater defensiveness; and the impersonal orientation has been positively related to self-derogation, poorer self-regulation, and depression.

Consistent with SDT, both the autonomy and controlled orientations were positively related to an internal locus of control (Rotter, 1966). The concept of internal locus of control concerns only whether one believes that outcomes are contingent on behavior, but it does not differentiate between whether the resulting motivation is autonomous or controlled. An internal perceived locus of causality, in contrast, reflects only autonomous motivation.

The causality orientations refer to relatively stable motivational orientations that parallel the three concepts of autonomy-supportive, [p. 421 ↓] controlling, and amotivating interpersonal contexts, and the causality orientations also parallel the state-like motivational concepts of autonomous motivation, controlled motivation, and amotivation. Importantly, research has shown that the causality orientations and the types of social contexts make parallel predictions for the three state-like motivations and also predict independent variance in the motivational states and in a range of other outcomes, such as work performance (Baard et al., 2004), and better maintained weightloss over two years (Williams, Ryan, and Deci, 1996). Causality orientations theory (COT) is more in the realm of personality than social psychology, but SDT concepts are all interrelated.

Organismic Integration Theory

The second question we were pondering in 1980, namely whether extrinsic motivation could become autonomous, led us to address the concept of *internalization* by differentiating between internalized extrinsic motivation that is controlled and internalized extrinsic motivation that is autonomous. The resulting conceptualization, with its manifold ramifications, was referred to as *organismic integration theory* (OIT; Deci and Ryan, 1985b; Ryan et al., 1985). It has at its core the assumption of an inherent integrative tendency, viewed as the fundamental developmental process, and this tendency, like intrinsic motivation, was predicted to be facilitated by support for the basic psychological needs and to be impaired by thwarting of the needs. To understand internalization, however, it became immediately clear that we needed the third basic psychological need, namely that of relatedness, in order to have a full understanding of internalization and integration. The *need for relatedness* – that is, the need to be close to, trusting of, caring for, and cared for by others – is similar to what Baumeister and Leary (1995) would later call the *need to belong*. Since 1985 we have worked with the concept of the three basic and universal psychological needs without having found a compelling reason to add a fourth.

Perhaps the most important and far-reaching element of OIT is its differentiation of the varied types of internalized extrinsic motivation. Whereas most internalization theories have considered a value or behavioral regulation to be either outside the person or inside the person (with inside being better), OIT specified varying degrees to which a

behavioral regulation and its accompanying value could be internalized, resulting in different types of subsequent regulation. A relatively unstable form of internalization is represented by *introjection*, in which people adopt an ambient value or practice, and are motivated to maintain it, as they “should,” in order to maintain self-approval or avoid guilt. Self-esteem contingencies (Deci and Ryan, 1995) and ego-involvement (Ryan 1982) are both forms of introjected regulation. A second type of internalization is *identification*, which involves personally identifying with the value of a behavior and thus fully accepting it as their own. A final type of internalization is *integration*, in which people integrate identifications with other aspects of their core values and practices. Internalization is a manifestation of the organismic integration process, and integrated regulation results when that process has worked to fully internalize a behavioral regulation.

By specifying the degrees of internalization OIT provided an account for the troubling issue of people forcing themselves to behave using internal pressures such as contingent self-esteem or threats of guilt. These processes are internal to people but they are by no means optimally healthy ways for people to regulate themselves for they do not possess the qualities of autonomy – namely, flexibility, volition, and a sense of choice. It turns out, in fact, that introjection has correlates and consequences that are qualitatively closer to external control than to identification (e.g., Ryan et al., 1993). In contrast to introjection, identification and integration share many qualities with intrinsic motivation and represent relatively autonomous forms of extrinsic motivation. These styles of regulation, together with *external regulation* [p. 422 ↓] (behavior controlled by external reward or punishment contingencies) and *intrinsic motivation*, represent five ways of regulating oneself, and we have proposed that the reasons people engage in a behavior can be aligned with the various types of motivation and regulation – namely, external reasons, introjected reasons, identified reasons, integrated reasons, and intrinsic reasons.

With this new conceptualization, the most salient and important distinction within SDT is neither “intrinsic versus extrinsic motivation” nor “internal versus external to the person,” but is rather *autonomous versus controlled motivation*. Autonomous motivation encompasses intrinsic motivation and identified/integrated extrinsic motivation, whereas controlled motivation comprises external control and introjected regulation. Moreover

the issue of autonomy was considered relative, insofar as most behaviors represent a mixture of the various reasons for acting described by these five categories.

OIT has been vigorously researched, frequently using an assessment approach developed by Ryan and Connell (1989) that measures the degree to which individuals do particular behaviors for various autonomous and controlled reasons. Using this approach, for example, researchers have found that the types of regulation form a *simplex pattern* indicating that they fall along a relative-autonomy continuum anchored by external regulation on the controlling end and integrated regulation and intrinsic motivation on the autonomous end. Further, studies have shown that the more autonomous types of motivation are associated with such outcomes as wellness, engagement in work or schoolwork, perceived competence, and deeper conceptual learning (e.g., Grolnick and Ryan, 1987; Ryan and Connell, 1989; Vallerand, 1997).

Internalization is particularly important in childhood, but it is relevant across the lifespan. For example, in our research on health behavior change among adults, we view the change process as being based in internalization of the value for and regulation of healthy behaviors (Williams et al., 1998). Research has shown that the more fully a regulation is internalized, such that the behavior is more autonomous, the more likely people are to change and maintain behaviors such as eating healthy diets and stopping smoking (e.g., Williams et al., 2006). Further, internalization plays a similarly important role in making positive changes within psychotherapy (Pelletier et al., 1997). This general formulation has been used in many life domains, as discussed later in the chapter.

Given the importance of well-internalized extrinsic motivation for effective functioning and well-being, we soon began examining the conditions most likely to promote full internalization. We hypothesized that social environments that support satisfaction of the basic psychological needs would promote fuller internalization of extrinsic motivation, and many studies have focused on the relation of autonomy support to internalization. *Autonomy support* refers to taking the others' perspective, encouraging initiation and exploration, providing choice, and being responsive to the others. In an interview study with parents, Grolnick and Ryan (1989) found that when parents were more autonomy supportive rather than controlling in their child rearing, the children more fully internalized the regulation for doing their schoolwork and chores around the

house. Further, a laboratory experiment by Deci et al. (1994) revealed that providing a meaningful rationale for engaging in an uninteresting task, acknowledging people's feelings about the task, and making requests in ways that are more consistent with choice than with control all contributed to greater internalization and integration.

One common way for parents to be controlling rather than autonomy supportive is through *parental conditional regard*. Parents provide more attention and affection when the children do or be what the parents want, and they withdraw attention and love when the children do not live up to the parents' expectations. Assor et al. (2004) showed that [p. 423 ↓] when parents are conditionally regarding, their children tend to introject the demands – thus conditionally administering esteem to themselves as their parents had done to them. In the process, they experience internal conflict, shame and guilt when they fail, and anger and resentment toward their parents. Follow-up research compared parental conditional regard with parental autonomy support (Roth et al., 2009). Results showed that conditional regard predicted introjection, both suppression of and dysregulation of emotions, and both controlled motivation and amotivation in school; whereas autonomy support led to the experience of choice, integrated regulation of emotions, and interest-focused school engagement. In short, the consequences of autonomy support are far more adaptive than those of conditional regard.

Although much OIT research has focused on autonomy support for facilitating internalization, the theory proposes that it is satisfaction of all three needs that is essential for full internalization. It turns out that parents and other authorities who support autonomy also tend to support competence and relatedness, so it is often the case that, when autonomy is being supported, competence and relatedness are also being supported, although satisfaction of each need is associated with independent influences and resulting dynamic outcomes.

In sum, research guided by OIT has shown that extrinsic motivation can be internalized to differing degrees, leading to types of internal regulation that differ in the degree to which they represent autonomy. The types that are more autonomous (identified and integrated) have been associated with more positive outcomes in various domains including schoolwork and regulation of emotions, whereas the types of regulation that are controlling (external and introjection) are associated with poorer outcomes

across domains. Finally, social contexts that support the basic needs for autonomy, competence, and relatedness also facilitate fuller internalization, whereas context that thwart need satisfaction, such as using rewards and punishments or conditional regard promote only introjection and are accompanied by indicators of ill-being.

OIT is perhaps best viewed as a theory of personality development and self-regulation, although the idea that social contexts facilitate versus impair internalization by supporting versus thwarting the basic psychological needs is very much a social-psychological concept. In this regard we see that the more the social contexts are relation-ally accepting and inclusive, competence promoting, and autonomy supportive, the more fully people internalize ambient social values and norms. Hence, social-contextual conditions that facilitate internalization of extrinsic motivation share much with those that maintain or enhance intrinsic motivation.

By the late 1980s, SDT's three interrelated mini-theories (CET, COT, and OIT) were being validated by considerable research. Yet, as that research was accumulating, we also saw that the concept of basic psychological needs had additional utility. In contexts where the three basic needs were being satisfied, people were more likely to exhibit both intrinsic motivation and more integrated forms of extrinsic motivation. But just as importantly, across these studies we also observed that where needs were satisfied, participants were invariably reporting greater wellness; whereas when any of the needs was thwarted, various forms of defensiveness and ill-being were evident. This led us to formulate the fourth mini-theory focused on the core definition of basic psychological needs for autonomy, competence, and relatedness and their roles as essential nutriments for healthy development, well-being, and mature relationships. We now take a brief look at this fourth mini-theory and some of the research it organizes.

Basic Psychological Needs Theory

The basic psychological needs theory (BPNT; Ryan et al., 1996) was built on the concept of universal psychological needs, which had been important since the earliest days of [p. 424 ↓] our work. BPNT was formulated primarily to account for the well-being effects associated with satisfactions of autonomy, competence, and relatedness needs, and much research has shown both the necessity of basic need satisfaction

for well-being and the importance of the basic needs as mediators of the effects of social contexts on wellness. For example, need satisfaction at the between-person level has predicted better performance and greater psychological health in the workplace (e.g., Baard et al., 2004), and increases in need satisfaction over time have mediated the longitudinal relations between autonomy support of law school professors and the well-being of law students (Sheldon and Krieger, 2007). Other BPNT research examined basic need satisfaction not only at the between-person level but also at the within-person level and found that general need satisfaction was associated with greater psychological health at the person level, and also that on a daily basis, people experienced greater positive affect and less negative affect on those days when they got more satisfaction of their basic needs (Reis et al., 2000; Ryan et al., 2010). As we will see later, much research has employed the concept of basic psychological needs across domains, contexts, and cultures.

As we were formulating BPNT, work by Kasser and Ryan (1993) had begun examining the importance of different goal contents. In that work, life goals were differentiated, based on factor analysis, as being either intrinsic (directly satisfying of the basic needs) or extrinsic (more distal from the needs and may be antagonistic to them). We initially interpreted the goal-content work with BPNT, but more recently we recognized that this research area had become so extensive and complex that it required its own mini-theory. The fifth mini-theory within SDT is thus referred to as goal content theory (GCT).

Goal Content Theory

Based on the factor analyses, intrinsic aspirations or life goals included personal growth, affiliation, and community, whereas extrinsic goals included wealth, fame, and image (Kasser and Ryan, 1996). The studies showed that, when people rated the extrinsic aspirations as being strong relative to the intrinsic aspirations, they evidenced less self-actualization and vitality, and more depression, anxiety, and narcissism. Whereas the line of work initiated by Kasser and Ryan looked at aspirations or life goals as individual differences, other studies by Vansteenkiste and colleagues manipulated the salience of people's goals and found that focusing people on extrinsic goals led to poorer performance on learning activities (e.g., Vansteenkiste et al., 2004). In short, strongly pursuing extrinsic life goals, whether because of an individual difference or a

prompt, has led to less well-being, more ill-being, and poorer performance, presumably because the extrinsic aspirations do not directly satisfy the basic needs, and indeed often crowd out or compromise their satisfaction. For example, as materialists spend effort and time accumulating, they often compromise autonomy and relatedness in the pursuit of more “stuff.” Additional research showed that not only pursuing extrinsic aspirations, but also attaining them, could be problematic for psychological health (Niemiec et al., 2009). Specifically, whereas attainment of intrinsic aspirations was associated with greater well-being and less ill-being, mediated by basic psychological need satisfaction, attainment of extrinsic aspirations did not enhance well-being but did relate to greater ill-being.

Research further showed that people develop stronger intrinsic life goals when their parents are accepting, affirming, and autonomy supportive (Williams et al., 2000), but they develop stronger extrinsic goals when their parents are rejecting and controlling. Presumably, when parents are cold and pressuring, their children fail to experience adequate need satisfaction, thus experiencing insecurity and developing what we call need substitutes, such as the pursuit of wealth, fame, and image. These goals then guide subsequent actions, which result in additional [p. 425 ↓] need thwarting and an ongoing vicious cycle.

The Continued Expansion of SDT

In recent years, SDT has continued to expand in order to incorporate a broad range of new research topics. We have not, however, specified new mini-theories beyond the five already discussed, but instead have used just the SDT macro-theory to guide and interpret this new work. In this section we briefly address four of the basic research topics that have been examined within SDT during the past decade or so.

Cross-Cultural Studies

SDT states strongly that the three basic psychological needs are universal such that their satisfaction versus thwarting affects the psychological well-being of all people. This proposition has two important implications. First, it requires that the proposition

be compatible with an evolutionary perspective, and second that the relation between need satisfaction-versus-thwarting and well-being be confirmed across a variety of cultures with different economic and political systems and with different cultural values. Elsewhere (Deci and Ryan, 2000), we presented an argument that the proposition of universal psychological needs for competence, relatedness, and autonomy is indeed consistent with an evolutionary perspective. Further, there have now been several studies supporting the cross-cultural relevance of the needs proposition. Here we briefly discuss two of those studies.

Within psychology, the idea of a need for relatedness (or belongingness or love) is fairly widely accepted and there is little argument about it being relevant across cultures. In addition, the idea of a need for competence is consistent with several prominent theories, and its relevance to multiple cultures has not been contested. On the other hand, however, the concept of a basic, universal psychological need for autonomy is very controversial, with cross-cultural theorists such as Markus et al. (1996) maintaining that people acquire their needs from their cultures and that East Asian cultures do not value autonomy and independence but instead value relatedness and interdependence. Hence, these authors maintain that the SDT concept of autonomy is relevant only to Western cultures where individualism is valued. This, of course, implies that autonomy is not a universal need.

From the SDT perspective, autonomy across cultures would be evident in two ways. First, the prototype of autonomy is the intrinsic motivation inherent to the nature of people, so it should be apparent regardless of culture. Second, autonomy should also be manifest in all cultures as behaviors that are motivated by well-internalized extrinsic motivations. Most of the cross-cultural research has focused on the second.

Still, anecdotally, it should be obvious to anyone who has observed young children in any culture that intrinsically motivated learning and play are ubiquitous. Of course, the degree of support for these activities and opportunities to enact intrinsically motivated behaviors may differ in different cultures, but the phenomenon of being self-motivated to move and laugh and play can be found in any place where it has not been stamped out. Further, there is some education research indicating that intrinsic motivation is affected by autonomy support versus control in the social context and that intrinsic motivation leads to more effective learning in Japan, China, Korea, and other collectivist contexts

(e.g., Bao and Lam, 2008; Jang et al., 2009; Kage and Namiki, 1990), just as is the case in the US.

In one cross-cultural study of autonomy, Chirkov et al. (2003) pointed out that autonomy within Western or Eastern cultures can accrue from relevant values having been fully internalized. For example, persons in an Eastern culture could be autonomous when enacting a collectivist cultural value, just as persons in a Western culture could be autonomous when enacting an individualist cultural [p. 426 ↓] value if they had fully internalized the target value. Chirkov et al. sampled students in Russia, Turkey, South Korea, and the US, who completed a self-regulation questionnaire assessing their reasons for engaging in various cultural practices. Results showed that the degree to which the participants had internalized (and expressed greater relative autonomy for enacting) the values and regulations for the various practices predicted their degree of psychological health and well-being. This confirmed that being autonomous was just as important for psychological wellness in Korea as in the US and each of the other countries in the study. Chirkov et al. found further that the relation between autonomy and well-being was not moderated by gender. This is important because theorists such as Jordan (1997) have argued that autonomy is a male characteristic and has little relevance to women's sensibilities. The Chirkov et al. study indicates that it was indeed relevant for both women and men in each culture studied. The researchers pointed out that part of the reason for the strongly disparate positions was that authors such as Markus et al. (1996) and Jordon tend to conflate autonomy (volition) and independence (nonreliance), rather than differentiating these important constructs as SDT research and theory have done (Deci and Ryan, 2000).

A study by Deci et al. (2001) examined working adults in the US, which has a capitalist economy, and Bulgaria, which at the time of data collection had a central planning economy for the majority of companies, which were still state-owned. The researchers found that the degree to which the workers in both countries perceived their managers to be autonomy supportive positively predicted satisfaction of the needs for autonomy, competence, and relatedness while at work. Need satisfaction, in turn, positively predicted engagement with work and psychological adjustment at work. Together, this study, the Chirkov et al. (2003) study, and many other more recent studies (e.g., Lynch et al., 2009) using varied methodologies have supported the view that basic psychological need satisfaction, and in particular satisfaction of the need for autonomy,

is important for psychological well-being across a range of cultures regardless of whether they value individualism or collectivism more strongly.

Close Personal Relationships

It is often stated that being in a meaningful relationship, such as with a romantic partner or best friend, requires that people relinquish autonomy to keep the relationship strong. The SDT perspective, however, is that feeling autonomous *within* the relationship is an essential element for the relationship to be strong and intimate. Consequently, SDT researchers have, over the past decade, engaged in research examining the importance of autonomy and autonomy support for the highest quality relationships.

Three studies by La Guardia et al. (2000) examined within-person variance in security of attachment among close relational partners (mother, father, best friend, and romantic partner). Attachment theorists suggest that children develop working models of attachment based largely on early interactions with primary caregivers, and these between-person individual differences strongly influence the security of attachment with all subsequent close attachment partners. The La Guardia et al. research indicated that there was between-person variance in security of attachment; however, there was also substantial variance explained at the within person level. The security of attachment that people had within their closest relationships varied substantially from partner to partner, and the felt security within each relationship was a function of the level of basic psychological need satisfaction, including autonomy, experienced within that relationship. Lynch et al. (2009) assessed autonomy support from close others in multiple cultural contexts and similarly found that relationship satisfaction as well as self-functioning were higher in relationships characterized by autonomy support. Patrick et al. (2007) found further that satisfaction of the basic psychological needs [p. 427 ↓] within romantic relationships predicted personal well-being, relational well-being, and effective management of conflict within the relationship. Yet another examination of autonomy in relationships focused on best friends (Deci et al., 2006). The researchers found that receiving autonomy support from a best friend was associated with greater relationship quality and greater well-being in the person receiving it. Further, the giving of autonomy support by that person to the best friend was also associated with the person experiencing greater relationship quality and well-being. In other words, both

receiving and giving autonomy support within a friendship accounted for significant independent variance in the person experiencing higher relationship quality and greater well-being.

From these and other studies, it seems clear that feeling a sense of autonomy and volition within close relationships is important for experiencing the relationships as satisfying (La Guardia and Patrick, 2008). Thus, feeling autonomy and relatedness are not inherently antagonistic but rather are mutually supportive, although it is possible to make these needs antagonistic, as when a relational partner provides conditional regard, or requires that the person give up his or her autonomy in order to receive affection or regard from the partner.

Vitality: Energy Available to the Self

Ryan and Frederick (1997) used the concept of subjective vitality to refer to the sense of aliveness and vigor that energizes volitional actions. They suggested that vitality results from satisfaction of the basic psychological needs, is an important indicator of health, and provides the necessary energy for effective self-regulation and coping with challenges. Ryan and Deci (2008) proposed that, whereas attempts to control oneself (i.e., to act in accord with one's introjects) can drain energy and diminish aliveness, autonomous self-regulation is not depleting but is instead vitalizing. Vitality and autonomous self-regulation are thus activating, but it is a type of activation involving positive affect and is different from the energy people experience when they are angry or anxious. Vitality, in short, is energy associated with people's integrated sense of self, which thus invigorates the processes of choice, volition, and effective coping with challenges (e.g., Rozanski, 2005).

Baumeister et al. (1998) proposed that any form of self-regulation consumes psychological energy so they predicted that having choice, which has been found in many situations to enhance autonomous motivation, should be depleting of energy and vitality. They reported results that they interpreted as supporting this assertion. However, in line with the SDT viewpoint, Moller et al. (2006) argued that true choice should not be depleting, and they pointed out that the condition Baumeister et al. had referred to as "high choice" actually represented a controlled rather than autonomous

condition (i.e., pressure toward a particular choice) and hence was not true choice. In contrast, Moller et al. (2006) found in three experiments, that when participants were given true choice with no pressure there was no depletion and it led to significantly more energy and vitality than the “controlled choice” condition used by Baumeister et al. (1998).

Other experiments have also found that controlled regulation (i.e., self-control) is depleting as SDT predicts but that autonomous regulation tended to be vitalizing (e.g., Muraven et al., 2008; Nix et al., 1999). In sum, SDT's distinction between controlled regulation (i.e., self-control) and autonomous regulation (i.e., true self-regulation) is critical for understanding both vitality and depletion. Self-control depletes energy and vitality, but self-regulation, which promotes psychological need satisfaction, is vitalizing.

Within SDT, vitality, which is the energy available to the self, has been linked to both physical health and satisfaction of psychological needs for competence, autonomy, and relatedness (Ryan and Frederick, 1997). Thus, although energy has diurnal patterns and biological underpinnings (e.g., Thayer, 2001) it also varies with the support versus thwarting [p. 428 ↓] of psychological needs affecting a wide range of outcomes from motivation to mood.

Nonconscious Processes and Mindful Awareness

Because we use “experience of choice” about one's behavior as an aspect of autonomy's definition, various writers have interpreted this to mean that autonomy necessitates conscious decision-making. However, a sense of choice does not require deliberate decision making, it merely requires endorsing one's actions. Thus, SDT allows for nonconscious initiation of autonomous behavior, and numerous experiments have now examined the nonconscious prompting of autonomous and controlled regulation of behavior.

Levesque and Pelletier (2003) primed autonomous and controlled motivation with words relevant to each concept and then allowed participants to spend 15 minutes

working on an interesting activity. Subsequent dependent measures indicated that those participants primed with autonomy-related words were more intrinsically motivated for the task than those primed with control-related words. Several experiments by Hodgins and colleagues have also shown that people primed with autonomy display less defensiveness than those primed with control. For example, Hodgins et al. (2006) used a word-priming procedure and then examined participants' self-handicapping around performance on a physical activity. Self-handicapping is a defensive response in which people deliberately do something (e.g., stay up very late the night before an important activity) to have an excuse in case they perform badly. The researchers found that people primed with control showed more self-handicapping than those primed with autonomy. These results, which were prompted with a priming procedure, parallel those found by Knee and Zuckerman (1998) in which self-reported control assessed with the causality orientations scale also led to more self-handicapping than did self-reported autonomy.

In short, nonconscious initiation of actions can be consistent with either autonomous or controlled functioning. Further, when people are lacking awareness they are more vulnerable to being controlled through primes; however, when controlled primes are present, people still have the capacity, through mindfulness, to be autonomous in spite of the primes. Since the early days of SDT (e.g. Deci and Ryan, 1980) we have proposed that *mindful awareness* facilitates more autonomous regulation of behavior, but beginning with Brown and Ryan (2003), the last few years have witnessed more vigorous research efforts linking mindfulness with more autonomous, self-endorsed regulation, less defensiveness, and thus greater need satisfaction and wellness.

Sdt Roots and Relations

The inspirations and foundations for SDT have come from several psychological traditions. The concept of intrinsic motivation first appeared in experimental psychology when Hullian drive theory concepts were found to be inadequate for explaining the exploratory behaviors of rats and monkeys (e.g., Harlow, 1950). White (1959) drew together that work and introduced the concept of competence as a fundamental motivation and basic need. Using these starting points, SDT has operated wholly within the empirical tradition since its beginning, and Heider's (1958) attribution theory

provided the route into the empirical work on intrinsic motivation at a time when the field of motivation was essentially moribund within psychology.

Although SDT has empirical roots in experimental social psychology (Heider, 1958), its theoretical roots extend farther afield to the organismic (Goldstein, 1939), ego-psychology (Hartmann, 1958), and existential-phenomenological (Pfander, 1910/1967) traditions, which focus on the critical importance of human experience and meaning in the determination of action, and the biologically inherent tendencies [p. 429 ↓] toward integrated functioning (see Ryan, 1995; Ryan and Deci, 2004). In other words, we share meta-theory and some aspects of theory with these latter traditions while working within the empirical tradition.

For example, when ego-psychologists within the psychoanalytic tradition abandoned psychosexual staging as its main theory of normal development, the idea of conflict-free ego energy emerged as being an inherent motivation rather than being a derivative of the id. This so-called independent ego energy (White, 1963) is evident in intrinsic motivation, and is a motivational underpinning of the developing ego and healthy self. Within SDT, intrinsic motivation, with its underlying psychological needs for autonomy, competence, and relatedness, energizes the operation of the organismic integration process. This integrative process, which we view as the inherent and natural developmental process, is energized by intrinsic motivation and involves the internalization and integration of attitudes, values, motivations, and emotional regulatory processes, and it has much in common with Loevinger's (1976) concept of ego-development, which is central to her structural stage theory of the ego and its regulation. Like Loevinger's theory, Piaget's (1971) theory of cognitive development is also an organismic theory for it too assumes an inherent developmental process toward assimilation and integration. SDT, although explicitly not a stage theory, has common threads with these and related organismic theories that assume a natural tendency toward development, which thus does not have to be "programmed" by the environment, although, as SDT has emphasized, environmental supports are necessary for effective functioning of the integrative process. Finally, humanistic theories (e.g., Rogers, 1963) also posit an inherent developmental process, which these theories tend to call self-actualization.

SDT does recognize age-related changes in motivation, but our focus is on: (1) the fundamental integrative process that is operative across the lifespan; (2) the basic psychological needs for autonomy, competence, and relatedness that energize the natural developmental process; and (3) the different regulatory processes, which, although they differ in the degree to which they reflect mature regulation, do not develop sequentially through age-related stages. Further, unlike stage theories, we maintain that adults are regulated to some degree by intrinsic motivation and each type of extrinsic motivation – external, introjected, identified, and integrated – and that people's motivational profiles will vary over time but not necessarily unidirectionally. Finally, although SDT is a theory of human needs and their relations with integrated functioning, it rejects Maslow's (1971) hierarchy of needs, instead asserting that the three basic needs are implicated across development (Ryan et al., 2006).

Perhaps SDT's major point of departure from these theories, other than that it being empirically based, is its focus on motivation. Whereas each of these other theories assumes inherent activity and an integrative tendency underlying development, they simply propose that the tendency functions without addressing the interaction between inherent needs and social conditions that supports its functioning. One of the most important reasons for addressing this issue is that it provides a means for predicting the conditions under which the developmental process will function most effectively. Specifically, it will function effectively to the degree that the needs for competence, autonomy, and relatedness are satisfied. This allows a theoretically based examination of the conditions that promote, for example, healthy child development, effective therapeutic change, optimal learning, skilled performance, and prosocial behavior.

Applications to Social Issues

SDT, perhaps as much as, or more than, any other theory of social psychology, has been applied to a broad range of life domains and [p. 430 ↓] social issues. Some of the studies have been longitudinal or cross-sectional field studies, some have been randomized trials, and some have been lab experiments. SDT, and specifically the CET mini-theory, has identified specific environmental events such as rewards, deadlines, threats of punishment, competition, and evaluations, as well as controlling interpersonal contexts, that tend to (1) undermine intrinsic motivation, (2) have a corrupting effect in

which people take the shortest path to the outcome, sometimes even when the path is inappropriate or immoral (Ryan and Brown, 2005), and (3) relate to poorer heuristic performance and well-being. SDT research has also identified external events such as choice, positive feedback, and acknowledging feelings, as well as autonomy-supportive social contexts, that enhance intrinsic motivation, internalization, and psychological well-being.

Because SDT research has shown that social contexts and communication styles affect motivation, performance, and well-being, many studies dealing with real-world contexts and social issues have examined autonomy-supportive versus controlling styles as they affect outcomes such as learning, socialization, healthy behaving, job satisfaction, prosocial behaviors, therapeutic outcomes, and rehabilitation in various treatment settings. We address just a small portion of this work, offering illustrative examples rather than a comprehensive review.

Promoting Healthy Behaving

Behavioral choices made by individuals everyday are among the most serious threats to physical health. For example, tobacco use has serious consequences such as heart disease and cancer; unhealthy diets and a lack of physical activity promote obesity, diabetes, and cardiovascular disease; and medication nonadherence works against the amelioration of illness. SDT research has tested process models for smoking cessation (Williams et al., 2002), weight loss (Williams et al., 1996), glucose management by diabetic patients (Williams et al., 2004), and medication adherence (Williams and Rodin et al., 1998), among other health issues, finding that autonomy support provided by physicians or other practitioners predicts patients' autonomous motivation and perceived competence, which in turn predicts maintained health behavior change, as well as concrete health indicators such as glycosylated hemoglobin or chemically verified cessation.

This consistent pattern of findings has led to clinical trials involving SDT-based interventions designed to be autonomy, competence, and relationally supportive. Randomized trials from our labs and others have thus far addressed tobacco cessation and abstinence, improved diet, exercise, and LDL cholesterol (Williams et al., 2006);

increased physical activity (Fortier et al., 2007); and better oral health (Halvari and Halvari, 2006). The Williams et al. trial found, for example, that an autonomy-supportive intervention led to significantly greater smoking cessation at the end of six months relative to that of a community-care control group, a significant difference that was still evident at both 18 and 30 months. Importantly, these patients were of relatively low socioeconomic status and over 50 percent of them said at their initial clinic visit that they did not want to make a quit attempt within the subsequent 30 days, belying the idea that a stage of “readiness” is a prerequisite for treatment and change.

Promoting Learning and Adjustment in Schools

SDT researchers have examined links among autonomy-support (relative to control) within classrooms and homes, students' autonomous motivation and perceived competence, and the outcomes of improved learning, achievement, and well-being. For example, Deci et al. (1981) found that autonomy-supportive classrooms led to increased intrinsic [p. 431 ↓] motivation and satisfaction of the competence need; Grolnick and Ryan (1987, 1989) found links from autonomy support to internalization of extrinsic motivation and well-being; Chirkov and Ryan (2001) found these relations in Russia; Reeve et al. (2002) found that a rationale for behaviors given in an autonomy-supportive way led to fuller internalization and more engagement in learning; both Benware and Deci (1984) and Grolnick and Ryan (1987) linked autonomy support to enhanced deep learning and conceptual understanding; and Vansteenkiste et al. (2004) found that intrinsic rather than extrinsic learning goals led to greater learning.

A network of findings that includes those just mentioned has led to the development of reform interventions that have used SDT as the primary theoretical basis (Deci, 2009). The SDT approach stands in clear contrast to the recent focus on incentives, accountability, and high-stakes testing within this country, where the assumption is that administrators, teachers, and students need to be held accountable for their performance (typically assessed as student scores on state-administered exams) and that the pressured accountability and the use of various incentives (i.e., high stakes) will motivate individuals at each level of the educational hierarchies to perform more

effectively. SDT's view is that the high stakes with the pressure surrounding them tend to undermine autonomous motivation for teaching and learning and promote various types of "gaming," perhaps the most extreme of which is "manipulating" tests scores and student records (see Ryan and Brown, 2005).

In contrast, school reform based on SDT has included work by Feinberg et al. (2007) in Israel in which they used an autonomy-supportive approach to teach the basic principles of SDT to administrators and teachers, and then facilitated a process of the school personnel creating and implementing strategies for improvement. In the US an approach to school reform developed by James Connell is a comprehensive structured approach with many SDT-related elements including improved teacher-student relationships by making smaller units within the schools, facilitating greater choice for teachers and students, and making instruction more optimally challenging and engaging. Results from evaluations of this approach have been very promising for improving attendance, graduation rates, and achievement (e.g., Gambone et al., 2004).

Psychotherapy and Behavior Change

In most psychotherapies, clients are encouraged to address and often change maladaptive behaviors, troubled relationships, or other presenting problems. These clients, in turn, have the choice about whether or not to engage in self-reflection and change. Successful psychotherapy requires a genuine willingness on the part of people to engage a process of change, particularly if the effects are expected to last beyond treatment (Deci and Ryan, 1985b; Ryan and Deci, 2008). SDT specifically argues that the maintenance and transfer of therapeutic changes, be they behavioral or psychological, requires the support of internalization and autonomous motivation (Pelletier et al., 1997). From the SDT perspective, this in turn suggests the importance of therapists providing an autonomy-supportive context for change.

SDT provides a broad, treatment-relevant framework for conceptualizing various developmental influences on psychopathology (Ryan et al., 2006). It also specifies the major elements of an approach to the practice of psychotherapy, arguing that the processes of growth and motivation entailed in psychological and behavioral change require an understanding of the dynamics of basic psychological needs and their

interpersonal support from practitioners (Ryan and Deci, 2008). SDT has been applied to problems from suicide (Britton et al., 2008) and depression (e.g., Zuroff et al., 2007), to drug and alcohol rehabilitation (e.g., Ryan et al., 1995; Zeldman et al., 2004). SDT, in fact, is not only an approach to psychotherapy but can [p. 432 ↓] also be applied to any existing interventions whose approach to motivating and implementing change processes affects clients' engagement and volition.

Sport and Physical Activity

One of the most vigorous applied areas of research within SDT is sport and physical activity. Studies have shown the role of both autonomous self-regulation and autonomy support in promoting greater motivation and persistence in physical activities, and SDT has been applied in physical education, health promotion, and coaching settings around the globe, as indicated by two recent collections of papers (Hagger and Chatzisarantis, 2007; Vlachopoulos, 2009).

Other Social Issues

Although SDT has relevance for many other social issues across a range of life domains, we very briefly mention just three more. The first concerns people experiencing job satisfaction and well-being in the workplace while performing effectively. For example, studies by Baard et al. (2004) in the banking industry found that employees who experienced more autonomy support from managers also reported more satisfaction of their basic psychological needs for autonomy, competence, and relatedness, and in turn had higher performance evaluations, higher well-being, and lower ill-being. Indeed, many studies point to the relevance of basic need supports for a productive, well-functioning workplace (e.g., Deci et al., 2001).

The second of these issue concerns prosocial behaviors. Recent studies by Weinstein and Ryan (2010) revealed that when participants were given choice (relative to no choice) over acts of helping, they in turn experienced significantly greater need satisfaction and displayed greater psychological well-being. Just as importantly these and other studies have shown that recipients of help benefit more, are less

psychologically threatened, and are more grateful when the help they receive is autonomously motivated.

Yet another intriguing area of research is motivation in virtual worlds. SDT research has demonstrated the role of need satisfaction in motivating videogame play, as well as the relation of basic needs to issues such as overuse and aggression in videogames, among others (e.g., Przybylski et al., 2009; Ryan et al., 2006).

As we stated, we could at best provide a very partial review of the expanding applied work based on SDT. In fact, SDT research is voluminous in areas of sustainability (Pelletier and Sharp, 2008), parenting (Grolnick, 2003), religion (Ryan et al., 1993), and the nature of happiness (Ryan and Deci, 2001), among others. We believe this is a result of philosophically well-founded and empirically testable theoretical foundations, as well as a focus on central issues in human experience and their impact on motivation and well-being.

Conclusion

SDT is a psychological macro-theory that focuses to a substantial extent on the effects of social-contextual factors on human motivation, behavior, and personality. In this chapter we explored the theory and its development, emphasizing the importance of autonomy-supportive, relative to controlling, interpersonal contexts for optimal motivation, effective behavior, healthy development, and psychological well-being. We presented each of the five mini-theories underlying SDT, as well as some newer areas of work that are extending that basic framework. We pointed out that SDT is an empirically derived theory, but its meta-theory has elements drawn from organismic, phenomenological, ego-psychological, and humanistic traditions, thus leading to basic assumptions and theoretical elements that make the theory [p. 433 ↓] very different from many other mainstream social psychological theories. Finally, we briefly reviewed applications of SDT that are relevant for addressing social issues, including health behavior change, education, psychotherapy, work motivation, virtual environments, and prosocial behavior.

References

- Assor, A., Roth, G. and Deci, E.L. (2004) The emotional costs of parents' conditional regard: A self-determination theory analysis . *Journal of Personality* , 72, 47–89. <http://dx.doi.org/10.1111/j.0022-3506.2004.00256.x>
- Bao, X. and Lam, S. (2008) Who makes the choice? rethinking the role of autonomy and relatedness in Chinese children's motivation . *Child Development* , 79, 269–283. <http://dx.doi.org/10.1111/j.1467-8624.2007.01125.x>
- Baard, P.P., Deci, E.L. and Ryan, R.M. (2004) Intrinsic need satisfaction: A motivational basis of performance and well-being in two work settings . *Journal of Applied Social Psychology* , 34, 2045–2068. <http://dx.doi.org/10.1111/j.1559-1816.2004.tb02690.x>
- Baumeister, R.F., Bratslavsky, E., Muraven, M. and Tice, D.M. (1998) Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology* , 74, 1252–1265. <http://dx.doi.org/10.1037/0022-3514.74.5.1252>
- Baumeister, R. and Leary, M.R. (1995) The need to belong: Desire for interpersonal attachments as a fundamental human motivation . *Psychological Bulletin* , 117, 497–529. <http://dx.doi.org/10.1037/0033-2909.117.3.497>
- Bem, D.J. (1972) Self-perception theory . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 6, 1–62. New York: Academic Press.
- Benware, C. and Deci, E.L. (1984) The quality of learning with an active versus passive motivational set . *American Educational Research Journal* , 21, 755–766.
- Britton, P.C., Williams, G.C. and Conner, K.R. (2008) Self-determination theory, motivational interviewing, and the treatment of clients with acute suicidal ideation . *Journal of Clinical Psychology* , 64, 52–66. <http://dx.doi.org/10.1002/jclp.20430>
- Brown, K.W. and Ryan, R.M. (2003) The benefits of being present: Mindfulness and its role in psychological well-being . *Journal of Personality and Social Psychology* , 84, 822–848. <http://dx.doi.org/10.1037/0022-3514.84.4.822>

Calder, B.J. and Staw, B.M. (1975) The interaction of intrinsic and extrinsic motivation: Some methodological notes . *Journal of Personality and Social Psychology* , 31, 76–80. <http://dx.doi.org/10.1037/h0076167>

Chirkov, V.I. and Ryan, R.M. (2001) Parent and teacher autonomy-support in Russian and U.S. adolescents: Common effects on well-being and academic motivation . *Journal of Cross Cultural Psychology* , 32, 618–635. <http://dx.doi.org/10.1177/0022022101032005006>

Chirkov, V.I., Ryan, R.M., Kim, Y. and Kaplan, U. (2003) Differentiating autonomy from individualism and independence: A self-determination theory perspective on internalization of cultural orientations and well-being . *Journal of Personality and Social Psychology* , 84, 97–110. <http://dx.doi.org/10.1037/0022-3514.84.1.97>

de Charms, R. (1968) Personal Causation . New York: Academic Press.

Deci, E.L. (1971). Effects of externally mediated rewards on intrinsic motivation . *Journal of Personality and Social Psychology* , 18, 105–115. <http://dx.doi.org/10.1037/h0030644>

Deci, E.L. (2009) Large-scale school reform as viewed from the self-determination theory perspective . *Theory and Research in Education* , 7, 247–255. <http://dx.doi.org/10.1177/1477878509104329>

Deci, E.L., Betley, G., Kahle, J., Abrams, L. and Porac, J. (1981) When trying to win: Competition and intrinsic motivation . *Personality and Social Psychology Bulletin* , 7, 79–83. <http://dx.doi.org/10.1177/014616728171012>

Deci, E.L. and Cascio, W.F. (1972) Changes in intrinsic motivation as a function of negative feedback and threats . Paper presented at the Eastern Psychological Association, Boston, April.

Deci, E.L., Connell, J.P. and Ryan, R.M. (1989) Self-determination in a work organization . *Journal of Applied Psychology* , 74, 580–590. <http://dx.doi.org/10.1037/0021-9010.74.4.580>

Deci, E.L., Eghrari, H., Patrick, B.C. and Leone, D.R. (1994) Facilitating internalization: The self-determination theory perspective . *Journal of Personality* , 62, 119–142. <http://dx.doi.org/10.1111/j.1467-6494.1994.tb00797.x>

Deci, E.L., Koestner, R. and Ryan, R.M. (1999) A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation . *Psychological Bulletin* , 125, 627–668. <http://dx.doi.org/10.1037/0033-2909.125.6.627>

Deci, E.L., La Guardia, J.G., Moller, A.C., Scheiner, M.J. and Ryan, R.M. (2006) On the benefits of giving as well as receiving autonomy support: Mutuality in close friendships . *Personality and Social Psychology Bulletin* , 32, 313–327. <http://dx.doi.org/10.1177/0146167205282148>

Deci, E.L. and Ryan, R.M. (1980). The empirical exploration of intrinsic motivational processes. In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 13, 39–80. New York: Academic Press.

Deci, E.L. and Ryan, R.M. (1985a) The General Causality Orientations Scale: Self-determination in personality . *Journal of Research in Personality* , 19, 109–134. <http://dx.doi.org/10.1016/0092-6566%2885%2990023-6>

Deci, E.L. and Ryan, R.M. (1985b) *Intrinsic Motivation and Self-determination in Human Behavior* . New York: Plenum.

Deci, E.L. and Ryan, R.M. (1995) Human autonomy: The basis for true self-esteem . In M. Kernis (ed.) *Efficacy, Agency, and Self-esteem* , pp. 31–49. New York: Plenum.

Deci, E.L. and Ryan, R.M. (2000) The ‘what’ and ‘why’ of goal pursuits: Human needs and the self-determination of behavior . *Psychological Inquiry* , 11, 227–268. http://dx.doi.org/10.1207/S15327965PLI1104_01

Deci, E.L., Ryan, R.M., Gagné, M., Leone, D.R., Usunov, J. and Kornazheva, B.P. (2001) Need satisfaction, motivation, and well-being in the work organizations of a former Eastern Bloc country . *Personality and Social Psychology Bulletin* , 27, 930–942. <http://dx.doi.org/10.1177/0146167201278002>

Deci, E.L., Schwartz, A.J., Sheinman, L. and Ryan, R.M. (1981) An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence . *Journal of Educational Psychology* , 73, 642–650. <http://dx.doi.org/10.1037/0022-0663.73.5.642>

Eisenberger, R. and Cameron, J. (1996) Detrimental effects of reward: Reality or myth? *American Psychologist* , 51, 1153–1166. <http://dx.doi.org/10.1037/0003-066X.51.11.1153>

Feinberg, O., Kaplan, H., Assor, A. and Kanat-Maymon, Y. (2007) The concept of 'internalization' (based on SDT) as a guide for a school reform program . Paper presented at the Third International Conference on Self-Determination Theory, Toronto, May.

Fortier, M.S., Sweet, S.N., O'Sullivan, T.L. and Williams, G.C. (2007) A self-determination process model of physical activity adoption in the context of a randomized controlled trial . *Psychology of Sport and Exercise* , 8, 741–757. <http://dx.doi.org/10.1016/j.psychsport.2006.10.006>

Gambone, M.A., Klem, A.M., Summers, J.A., Akey, T.A. and Sipe, C.L. (2004) Turning the Tide: The Achievements of the First Things First Education Reform in the Kansas City, Kansas Public School District . Philadelphia, PA: Youth Development Strategies, Inc.

Goldstein, K. (1939) *The Organism* . New York: American Book Co.

Grolnick, W.S. (2003) *The psychology of parental control* . Mahwah, NJ: Erlbaum.

Grolnick, W.S. and Ryan, R.M. (1987) Autonomy in children's learning: An experimental and individual difference investigation . *Journal of Personality and Social Psychology* , 52, 890–898. <http://dx.doi.org/10.1037/0022-3514.52.5.890>

Grolnick, W.S. and Ryan, R.M. (1989) Parent styles associated with children's self-regulation and competence in school . *Journal of Educational Psychology* , 81, 143–154. <http://dx.doi.org/10.1037/0022-0663.81.2.143>

Hagger, M.S. and Chatzisarantis, N.L. (2007) Intrinsic Motivation and Self-determination in Exercise and Sport . Champaign, IL: Human Kinetics.

Halvari, A.M. and Halvari, H. (2006) Motivational predictors of change in oral health: An experimental test of self-determination theory . Motivation and Emotion , 30, 295–306. <http://dx.doi.org/10.1007/s11031-006-9035-8>

Harlow, H.F. (1950) Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys . Journal of Comparative and Physiological Psychology , 43, 289–294. <http://dx.doi.org/10.1037/h0058114>

Hartmann, H. (1958) Ego Psychology and the Problem of Adaptation . New York: International Universities Press. (Originally published 1939.) <http://dx.doi.org/10.1037/13180-000>

Heider, F. (1958) The Psychology of Interpersonal Relations . New York: Wiley. <http://dx.doi.org/10.1037/10628-000>

Hodgins, H.S., Yacko, H.A. and Gottlieb, E. (2006) Autonomy and nondefensiveness . Motivation and Emotion , 30, 283–293. <http://dx.doi.org/10.1007/s11031-006-9036-7>

Jang, H., Reeve, J., Ryan, R.M. and Kim, A. (2009) Can self-determination theory explain what underlies the productive, satisfying learning experiences of collectivistically oriented Korean students? Journal of Educational Psychology , 101, 644–661. <http://dx.doi.org/10.1037/a0014241>

Jordan, J.V. (1997) Do you believe that the concepts of self and autonomy are useful in understanding women? In J.V. Jordan (ed.), Women's Growth in Diversity: More Writings from the Stone Center , pp. 29–32. New York: Guilford Press.

Kage, M. and Namiki, H. (1990) The effects of evaluation structure on children's intrinsic motivation and learning . Japanese Journal of Educational Psychology , 38, 36–45.

Kasser, T. and Ryan, R.M. (1993) A dark side of the American dream: Correlates of financial success as a central life aspiration . Journal of Personality and Social Psychology , 65, 410–422. <http://dx.doi.org/10.1037/0022-3514.65.2.410>

Kasser, T. and Ryan, R.M. (1996) Further examining the American dream: Differential correlates of intrinsic and extrinsic goals . *Personality and Social Psychology Bulletin* 22, 80–87. <http://dx.doi.org/10.1177/0146167296223006>

Knee, C.R. and Zuckerman, M. (1998) A nondefensive personality: Autonomy and control as moderators of defensive coping and self-handicapping . *Journal of Research in Personality* , 32, 115–130. <http://dx.doi.org/10.1006/jrpe.1997.2207>

Koestner, R., Ryan, R.M., Bernieri, F. and Holt, K. (1984) Setting limits on children's behavior: The differential effects of controlling versus informational styles on intrinsic motivation and creativity . *Journal of Personality* , 52, 233–248. <http://dx.doi.org/10.1111/j.1467-6494.1984.tb00879.x>

La Guardia, J.G. and Patrick, H. (2008) Self-determination theory as a fundamental theory of close relationships . *Canadian Psychology* , 49, 201–209. <http://dx.doi.org/10.1037/a0012760>

La Guardia, J.G., Ryan, R.M., Couchman, C. E. and Deci, E.L. (2000) Within-person variation in security of attachment: A self-determination theory perspective on attachment, need fulfillment, and well-being . *Journal of Personality and Social Psychology* , 79, 367–384. <http://dx.doi.org/10.1037/0022-3514.79.3.367>

Lepper, M.R., Greene, D. and Nisbett, R.E. (1973) Undermining children's intrinsic interest with extrinsic rewards: A test of the 'overjustification' hypothesis . *Journal of Personality and Social Psychology* , 28, 129–137. <http://dx.doi.org/10.1037/h0035519>

Levesque, C. and Pelletier, L.G. (2003) On the investigation of primed and chronic autonomous and heteronomous motivational orientations . *Personality and Social Psychology Bulletin* , 29, 1570–1584. <http://dx.doi.org/10.1177/0146167203256877>

Loevinger, J. (1976) *Ego Development* . San Francisco: Jossey-Bass.

Lynch, M.F., La Guardia, J.G. and Ryan, R.M. (2009) On being yourself in different cultures: Ideal and actual self-concept, autonomy support, and well-being in China, Russia, and the United States . *Journal of Positive Psychology* , 4, 290–394. <http://dx.doi.org/10.1080/17439760902933765>

Markus, H.R., Kitayama, S. and Heiman, R.J. (1996) Culture and basic psychological principles . In E.T. Higgins, ed. and A.W. Kruglanski (eds), *Social Psychology: Handbook of Basic Principles* , pp. 857–913. New York: Guilford Press.

Maslow, A.H. (1971) *The Farther Reaches of Human Nature* . New York: Viking Press.

Moller, A.C., Deci, E.L. and Ryan, R.M. (2006) Choice and ego-depletion: The moderating role of autonomy . *Personality and Social Psychology Bulletin* , 32, 1024–1036. <http://dx.doi.org/10.1177/0146167206288008>

Muraven, M., Gagné, M. and Rosman, H. (2008) Helpful self-control: Autonomy support, vitality, and depletion . *Journal of Experimental Social Psychology* , 44, 573–585. <http://dx.doi.org/10.1016/j.jesp.2007.10.008>

Niemiec, C.P., Ryan, R.M. and Deci, E.L. (2009) The path taken: Consequences of attaining intrinsic and extrinsic aspirations in post-college life . *Journal of Research in Personality* , 43, 291–306. <http://dx.doi.org/10.1016/j.jrp.2008.09.001>

Nix, G.A., Ryan, R.M., Manly, J.B. and Deci, E.L. (1999) Revitalization through self-regulation: The effects of autonomous and controlled motivation on happiness and vitality . *Journal of Experimental Social Psychology* , 35, 266–284. <http://dx.doi.org/10.1006/jesp.1999.1382>

Patrick, H., Knee, C.R., Canevello, A. and Lonsbury, C. (2007) The role of need fulfillment in relationship functioning and well-being: A self-determination theory perspective . *Journal of Personality and Social Psychology* , 92, 434–457. <http://dx.doi.org/10.1037/0022-3514.92.3.434>

Pelletier, L.G. and Sharp, E. (2008) Persuasive communication and proenvironmental behaviours: How message tailoring and message framing can improve integration of behaviors through self-determined motivation . *Canadian Psychology* , 49, 210–217. <http://dx.doi.org/10.1037/a0012755>

Pelletier, L.G., Tuson, K.M. and Haddad, N.K. (1997) Client motivation for therapy scale: A measure of intrinsic motivation, extrinsic motivation, and amotivation for

therapy . *Journal of Personality Assessment* , 68, 414–435. http://dx.doi.org/10.1207/s15327752jpa6802_11

Pfander, A. (1910/1967) *Phenomenology of Willing and Motivation* . Chicago: Northwestern.

Piaget, J. (1971) *Biology and Knowledge* . Chicago: University of Chicago Press.

Pritchard, R.D., Campbell, K.M. and Campbell, D.J. (1977) Effects of extrinsic financial rewards on intrinsic motivation . *Journal of Applied Psychology* , 62, 9–15. <http://dx.doi.org/10.1037/0021-9010.62.1.9>

Przybylski, A.K., Ryan, R.M. and Rigby, C.S. (2009) The motivating role of violence in video games . *Personality and Social Psychology Bulletin* , 35, 243–259. <http://dx.doi.org/10.1177/0146167208327216>

Reeve, J., Jang, H., Hardre, P. and Omura, M. (2002) Providing a rationale in an autonomy-supportive way as a strategy to motivate others during an uninteresting activity . *Motivation and Emotion* , 26, 183–207. <http://dx.doi.org/10.1023/A:1021711629417>

Reis, H.T., Sheldon, K.M., Gable, S.L., Roscoe, J. and Ryan, R.M. (2000) Daily well-being: The role of autonomy, competence, and relatedness . *Personality and Social Psychology Bulletin* , 26, 419–435. <http://dx.doi.org/10.1177/0146167200266002>

Rogers, C. (1963) The actualizing tendency in relation to 'motives' and to consciousness . In M.R. Jones (ed.), *Nebraska Symposium on Motivation* , 11, 1–24. Lincoln, NE: University of Nebraska Press.

Roth, G., Assor, A., Niemiec, C.P., Ryan, R.M. and Deci, E.L. (2009) The emotional and academic consequences of parental conditional regard: Comparing conditional positive regard conditional negative regard and autonomy support as parenting practices . *Developmental Psychology* , 45, 1119–1142. <http://dx.doi.org/10.1037/a0015272>

Rotter, J. (1966) Generalized expectancies for internal versus external control of reinforcement . *Psychological Monographs* , 80(1, Whole No. 609), 1–28.

Rozanski, A. (2005) Integrating psychologic approaches into the behavioral management of cardiac patients . *Psychosomatic Medicine* , 67, S67–S73.

Ryan, R.M. (1982) Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory . *Journal of Personality and Social Psychology* , 43, 450–461. <http://dx.doi.org/10.1037/0022-3514.43.3.450>

Ryan, R.M. (1995) Psychological needs and the facilitation of integrative processes . *Journal of Personality* , 63, 391–421. <http://dx.doi.org/10.1111/j.1467-6494.1995.tb00501.x>

Ryan, R.M., Bernstein, J.H. and Brown, K.W. (2010) Weekends, work, and well-being: Psychological need satisfactions and day of the week effects on mood, vitality, and physical symptoms . *Journal of Social and Clinical Psychology* , 29, 95–122. <http://dx.doi.org/10.1521/jscp.2010.29.1.95>

Ryan, R.M. and Brown, K.W. (2005) Legislating competence: The motivational impact of high stakes testing as an educational reform . In C. Dweck, ed. and A.E. Elliot (eds), *Handbook of Competence* , pp. 354–374. New York: Guilford Press.

Ryan, R.M. and Connell, J.P. (1989) Perceived locus of causality and internalization: Examining reasons for acting in two domains . *Journal of Personality and Social Psychology* , 57, 749–761. <http://dx.doi.org/10.1037/0022-3514.57.5.749>

Ryan, R.M., Connell, J.P. and Deci, E.L. (1985) A motivational analysis of self-determination and self-regulation in education . In C. Ames, ed. and R.E. Ames (eds), *Research on Motivation in Education: The Classroom Milieu* , pp. 13–51. New York: Academic Press.

Ryan, R.M. and Deci, E.L. (2001) On happiness and human potentials: A review of research on hedonic and eudaimonic well-being . In S. Fiske (ed.) *Annual Review of Psychology* , 52, 141–166. Palo Alto, CA: Annual Reviews, Inc.

Ryan, R.M. and Deci, E.L. (2004) Autonomy is no illusion: Self-determination theory and the empirical study of authenticity, awareness, and will . In J. Greenberg, ed. ,

S.L. Koole, ed. , and T. Pyszczynski (eds), *Handbook of Experimental Existential Psychology* , pp. 449–479. New York: Guilford Press.

Ryan, R.M. and Deci, E.L. (2008) From ego depletion to vitality: Theory and findings concerning the facilitation of energy available to the self . *Social and Personality Psychology Compass* , 2, 702–717. <http://dx.doi.org/10.1111/j.1751-9004.2008.00098.x>

Ryan, R.M., Deci, E.L., Grolnick, W.S. and La Guardia, J.G. (2006) The significance of autonomy and autonomy support in psychological development and psychopathology . In D. Cicchetti, ed. and D.J. Cohen (eds), *Developmental Psychopathology*, Vol 1: Theory and Method , 2nd Edition, pp. 795–849. New Jersey: Wiley.

Ryan, R.M. and Frederick, C.M. (1997) On energy, personality, and health: Subjective vitality as a dynamic reflection of well-being . *Journal of Personality* , 65, 529–565. <http://dx.doi.org/10.1111/j.1467-6494.1997.tb00326.x>

Ryan, R.M., Mims, V. and Koestner, R. (1983) Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory . *Journal of Personality and Social Psychology* , 45, 736–750. <http://dx.doi.org/10.1037/0022-3514.45.4.736>

Ryan, R.M., Plant, R.W. and O'Malley, S. (1995) Initial motivations for alcohol treatment: Relations with patient characteristics, treatment involvement and dropout . *Addictive Behaviors* , 20, 279–297. <http://dx.doi.org/10.1016/0306-4603%2894%2900072-7>

Ryan, R.M., Rigby, S. and King, K. (1993) Two types of religious internalization and their relations to religious orientations and mental health . *Journal of Personality and Social Psychology* , 65, 586–596. <http://dx.doi.org/10.1037/0022-3514.65.3.586>

Ryan, R.M., Rigby, C.S. and Przybylski, A. (2006) Motivation pull of video games: A self-determination theory approach . *Motivation and Emotion* , 30, 347–365.

Ryan, R.M., Sheldon, K.M., Kasser, T. and Deci, E.L. (1996) All goals are not created equal: An organismic perspective on the nature of goals and their regulation . In P.M. Gollwitzer, ed. and J.A. Bargh (eds), *The Psychology of Action: Linking Cognition and Motivation to Behavior* , pp. 7–26. New York: Guilford.

Scott, W.E., Jr. (1975) The effects of extrinsic rewards on 'intrinsic motivation': A critique . *Organizational Behavior and Human Performance* , 15, 117–129. <http://dx.doi.org/10.1016/0030-5073%2876%2990032-5>

Sheldon, K.M. and Krieger, L.S. (2007) Understanding the negative effects of legal education on law students: A longitudinal test of self-determination theory . *Personality and Social Psychology Bulletin* , 33, 883–897. <http://dx.doi.org/10.1177/0146167207301014>

Thayer, R.E. (2001) Calm Energy: How People Regulate Mood with Food and Exercise . New York: Oxford University Press.

Tooby, J. and Cosmides, L. (1992) The psychological foundations of culture . In J.H. Barkow, ed. , L. Cosmides, ed. , and J. Tooby, ed. , *The Adapted Mind* , pp. 19–136. New York: Oxford University Press.

Vallerand, R.J. (1997) Toward a hierarchical model of intrinsic and extrinsic motivation . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 29, 271–360. San Diego: Academic Press.

Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K.M. and Deci, E.L. (2004) Motivating learning, performance, and persistence: The synergistic effects of intrinsic goal contents and autonomy-supportive contexts . *Journal of Personality and Social Psychology* , 87, 246–260. <http://dx.doi.org/10.1037/0022-3514.87.2.246>

Vlachopoulos, S.P. (2009) Prologue . *Hellenic Journal of Psychology* , 6, vii-x.

Weinstein, N. and Ryan, R.M. (2010) When helping helps: Autonomous motivation for prosocial behavior and its influence on well-being for the helper and recipient . *Journal of Personality and Social Psychology* , 98, 222–244. <http://dx.doi.org/10.1037/a0016984>

White, R.W. (1959) Motivation reconsidered: The concept of competence . *Psychological Review* , 66, 297–333. <http://dx.doi.org/10.1037/h0040934>

White, R.W. (1963) Ego and Reality in Psychoanalytic Theory . New York: International Universities Press.

Williams, G.C., Cox, E.M., Hedberg, V. and Deci, E.L. (2000) Extrinsic life goals and health risk behaviors in adolescents . *Journal of Applied Social Psychology* , 30, 1756–1771. <http://dx.doi.org/10.1111/j.1559-1816.2000.tb02466.x>

Williams, G.C., Deci, E.L. and Ryan, R.M. (1998) Building health-care partnerships by supporting autonomy . In A.L. Suchman, ed. , P. Hinton-Walker, ed. , and R. Botelho (eds), *Partnerships in Healthcare* , pp. 67–87. Rochester, NY: University of Rochester Press.

Williams, G.C., Gagné, M., Ryan, R.M. and Deci, E.L. (2002) Facilitating autonomous motivation for smoking cessation . *Health Psychology* , 21, 40–50. <http://dx.doi.org/10.1037/0278-6133.21.1.40>

Williams, G.C., Grow, V.M., Freedman, Z., Ryan, R.M. and Deci, E.L. (1996) Motivational predictors of weight loss and weight-loss maintenance . *Journal of Personality and Social Psychology* , 70, 115–126. <http://dx.doi.org/10.1037/0022-3514.70.1.115>

Williams, G.C., McGregor, H.A., Sharp, D., Kouides, R.W., Levesque, C., Ryan, R.M. and Deci, E.L. (2006) A self-determination multiple risk intervention trial to improve smokers' health . *Journal of General Internal Medicine* , 21, 1288–1294. <http://dx.doi.org/10.1111/j.1525-1497.2006.00621.x>

Williams, G.C., McGregor, H.A., Zeldman, A., Freedman, Z.R. and Deci, E.L. (2004) Testing a self-determination theory process model for promoting glycemic control through diabetes self-management . *Health Psychology* , 23, 58–66. <http://dx.doi.org/10.1037/0278-6133.23.1.58>

Williams, G.C., Rodin, G.C., Ryan, R.M., Grodnick, W.S. and Deci, E.L. (1998) Autonomous regulation and adherence to long-term medical regimens in adult outpatients . *Health Psychology* , 17, 269–276. <http://dx.doi.org/10.1037/0278-6133.17.3.269>

Zeldman, A., Ryan, R.M. and Fiscella, K. (2004) Client motivation, autonomy support and entity beliefs: Their role in methadone maintenance treatment . *Journal of Social and Clinical Psychology* , 23, 675–696. <http://dx.doi.org/10.1521/jscp.23.5.675.50744>

Zuckerman, M., Porac, J., Lathin, D., Smith, R. and Deci, E.L. (1978) On the importance of self-determination for intrinsically motivated behavior . Personality and Social Psychology Bulletin , 4, 443–446. <http://dx.doi.org/10.1177/014616727800400317>

Zuroff, D.C., Koestner, R., Moskowitz, D.S., McBride, C., Marshall, M. and Bagby, M. (2007) Autonomous motivation for therapy: A new common factor in brief treatments for depression . Psychotherapy Research , 17, 137–147. <http://dx.doi.org/10.1080/10503300600919380>

<http://dx.doi.org/10.4135/9781446249215.n21>

Handbook of Theories of Social Psychology: Volume 1

The Theory of Planned Behavior

Contributors: Icek Ajzen

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "The Theory of Planned Behavior"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n22>

Print pages: 438-460

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n22>

[p. 438 ↓]

Chapter 21: The Theory of Planned Behavior

Icek Ajzen, ed.

Abstract

This chapter describes the theory of planned behavior (TPB), a prominent reasoned action model, its conceptual foundation, its intellectual history, and the research it has generated. From its roots in propositional control and expectancy theory, the TPB emerged as a major framework for understanding, predicting, and changing human social behavior. According to the theory, intention is the immediate antecedent of behavior and is itself a function of attitude toward the behavior, subjective norm, and perceived behavioral control; and these determinants follow, respectively, from beliefs about the behavior's likely consequences, about normative expectations of important others, and about the presence of factors that control behavioral performance. Empirical support for the theory comes from a host of correlational studies demonstrating its ability to predict intentions and behavior as well as from interventions showing that changes in behavioral, normative, and control beliefs can produce changes in intentions, and that these changes in intentions are reflected in subsequent behavior. The chapter also considers the TPB's reasoned action approach in the context of recent work on automatic, nonconscious processes in human social behavior. It is argued that insight into automaticity can complement the understanding of behavior provided by a reasoned action approach.

Introduction

The tenets of a reasoned action approach to human behavior strike a familiar chord. We are introspectively aware of the thoughts and feelings that lead up to our decisions and we find in these processes a convincing explanation for our behavior. By this account, the immediate causes of human social behavior are neither mysterious nor outside conscious awareness. Behavior is performed not automatically or mindlessly but follows reasonably and consistently from the behavior-relevant information available to us. Ever since I entered graduate school at the University of Illinois in 1966 and began to work with Martin Fishbein, this reasoned action assumption has guided my theoretical approach and empirical research.

[p. 439 ↓] Many contemporary models of human social behavior bear the hallmarks of a reasoned action approach. Among these models are Bandura's (1986, 1997) social cognitive theory, Triandis's (1972) theory of subjective culture and interpersonal relations, the health belief model (Rosenstock et al., 1994), goal setting theory (Locke and Latham, 1994), the information-motivation-behavioral skills model (Fisher and Fisher, 1992), and the technology acceptance model (Davis et al., 1989). The discussion in this chapter, however, focuses on what is perhaps the most influential reasoned action approach, the theory of reasoned action (TRA), which I developed in close collaboration with Martin Fishbein (Ajzen and Fishbein, 1980; Fishbein and Ajzen, 1975) and my extension of this model, the TPB (Ajzen, 1985, 1991a, 2005a).

The roots of the TRA can be traced, at least in part, to a confrontation with radical behaviorism and its law of effect. According to operant conditioning principles, behaviors followed by rewarding events are reinforced and behaviors followed by punishing events are weakened. This process is assumed to be automatic, requiring neither conscious awareness of behavior-outcome contingencies nor any other higher-order cognitive mediation. Yet there is scant evidence in human adults for operant conditioning of this kind or, for that matter, for classical conditioning without awareness (see Brewer, 1974).

At the University of Illinois I was exposed to, and greatly influenced by, the ideas of Don Dulany who was working in the area of verbal learning. Dulany (1962, 1968) was

struck by the stark contrast between the behaviorist explanation of behavior and by our intuitive understanding of its determinants. In a series of laboratory experiments, Dulany set out to test the contrasting views. Participants in these experiments were shown pairs of sentences, one pair on each trial, and were asked to choose one sentence and read it aloud. They performed this task while seated in a chamber that was kept at a constant temperature of 110 degrees Fahrenheit and 35 percent humidity. Sentences containing certain words were followed, with predetermined frequencies, by either a hot stream of air across the participant's face, a cool stream of air, or a stream of air at chamber temperature.

Now, according to a strict interpretation of operant conditioning principles, the hot blast of air – being a punishing event – should reduce the likelihood of the verbal responses that preceded it; that is, a hot blast of air should reduce the likelihood that participants will, on subsequent trials, choose sentences containing the stimulus words that were followed by the punishing event. Conversely, the cool blast of air should serve to reinforce emission of the preceding type of sentences. Dulany, however, believed that the affectively positive or negative events did not strengthen or weaken response tendencies directly. Instead, he proposed that their effects were mediated by higher mental processes, specifically by the participant's interpretation of the events' significance. To test this idea, he manipulated – orthogonally to the cool, neutral, or hot stream of air – instructions about the event's significance. Some participants were told that the stream of air indicated a correct response, some that it indicated neither a correct nor an incorrect response, and some that it indicated an incorrect response.

Dulany assumed that, in this situation, participants could form two hypotheses. First, they could come to believe that certain verbal responses (sentence choices) were followed by a certain event (a cool, neutral, or hot stream of air) with some degree of probability. He termed this belief the *hypothesis of the distribution of reinforcement*. Second, participants could form the belief that the stream of air meant that they had just done what they were supposed to do, what they were not supposed to do, or what they were neither supposed to do nor avoid doing. He termed this belief the *hypothesis of the significance of a reinforcer*. These two hypotheses provide the foundation for [p. 440 ↓] Dulany's *theory of propositional control*. According to the theory, people form a conscious intention to select a certain response, and it is this behavioral intention (*B_I*) that determines the response actually emitted. The intention itself is a function of two

factors. The first is the hypothesis of the distribution of reinforcement (RHd) weighted by the subjective value of the reinforcer (RSv). The more strongly people believe that a certain response will lead to a certain outcome, and the more positively they value that outcome, the stronger their intention to produce the response in question. The second factor influencing intentions is the hypothesis of the distribution of reinforcement weighted by the hypothesis of the reinforcer's significance. This compound variable is termed the *behavioral hypothesis (BH)*. The behavioral hypothesis is a function of the extent to which a response is perceived to produce a certain outcome and the extent to which this outcome is believed to indicate a correct response, that is, a response that is expected of them. Intentions to produce the expected response are strengthened to the extent that participants are *motivated to comply (MC)* with what they think they are supposed to do. Dulany's (1968) theory of propositional control in relation to behavioral intentions is shown symbolically in Equation (21.1); the actual response is assumed to be a direct function of the behavioral intention.

$$BI = (RHd) (RSv) + (BH) (MC) \quad (21.1)$$

Dulany's theory does not, of course, imply that reinforcement is irrelevant; it only suggests that the effect of reinforcement on behavior is mediated by the beliefs people form about reinforcement contingencies and what these contingencies signify. Consistent with the theory, intentions to select sentences of certain kinds had a correlation of 0.94 with the actual sentences selected, and behavioral intentions were predicted with a multiple correlation of 0.88 from (RHd) (RSv) and (BH) (MC). Interestingly, instructions regarding the significance of the reinforcement had a much stronger effect on intentions and actual response selection than had the affective value of the reinforcement. In other words, participants were guided in their response selection primarily by what sentences they thought the experimenter wanted them to read aloud, rather than by the affective consequences of their choices. Thus, they were willing to endure a hot stream of air to their faces while seated in a hot environment if they thought that this was expected of them in the experiment (and if they were motivated to comply).

Theory of Reasoned Action

To me, the results of Dulany's experiments were quite compelling and difficult to reconcile with automatic strengthening of responses by response-contingent events. His work suggested instead that human behavior is mediated and indeed controlled by higher mental processes. Dulany (1968) attributed a causal or instrumental role to conscious, volitional processes, although he recognized that, with practice, voluntary responses will tend to habituate and conscious rules may become unconscious. The causal role of conscious awareness and volition in human social behavior remains controversial however (see Bargh and Chartrand, 1999; Wegner and Wheatley, 1999); I will return to this topic later in this chapter.

Whether always consciously experienced or not, the cognitive processes identified in Dulany's theory of propositional control have their counterparts in the TRA (Ajzen and Fishbein, 1980; Fishbein, 1967a; Fishbein and Ajzen, 1975), the forerunner of the TPB (Ajzen, 1991b). Consider, first, the hypothesis of the distribution of reinforcement. In the TRA, this hypothesis is termed a *behavioral belief*. It is defined as a person's subjective probability that performing a certain behavior will produce a particular outcome, and the subjective value of the reinforcer is designated the person's evaluation of that outcome. [p. 441 ↓] Of course, most behaviors of interest to social psychologists are capable of producing more than one outcome and therefore, in the TRA, people are assumed to hold multiple behavioral beliefs each of which links performance of the behavior to a different outcome. A study on alcohol and drug use among college students (Armitage et al., 1999: 306, Table 2) provides a concrete example. In this study, the following behavioral beliefs about using alcohol and marijuana were held with relatively high frequency: "Makes me more sociable," "Leads to me having poorer physical health," "Will result in my becoming dependent on it," "Will result in me getting into trouble with authority," and "Makes me feel good."

The second component in Dulany's model, the behavioral hypothesis, is termed a *normative belief* in the TRA. It is defined as a person's subjective probability that a particular normative referent (the experimenter in Dulany's case) wants the person to perform a given behavior. As in Dulany's model, this normative belief is weighted (multiplied) by the person's motivation to comply with the referent's perceived

expectation. However, in the TRA it is assumed that people can hold normative beliefs with respect to more than one referent individual or group. Commonly identified referents are a person's spouse or partner, close family, friends, and, depending on the behavior under consideration, coworkers, health professionals, and law enforcement authorities.

Attitude toward a Behavior: The Expectancy-Value Model

As noted above, people generally hold a number of behavioral beliefs in relation to any given behavior. Each of these beliefs links the behavior to an outcome, and each outcome has a certain subjective value. In the TRA it is assumed that these behavioral beliefs and outcome evaluations combine to produce an overall positive or negative *attitude toward the behavior*. Specifically, the subjective value or evaluation of each outcome contributes to the attitude in direct proportion to the person's subjective probability that the behavior produces the outcome in question. This *expectancy-value* model of attitude is shown in Equation (21.2), where A represents attitude toward a behavior, b_i

i

is the subjective probability or belief that the behavior produces outcome i , e_i

is the evaluation of outcome i , and the sum is over the total number of behavioral beliefs.

$$A \approx \sum b_i e_i \quad (21.2)$$

The multiplicative combination of subjective probabilities (beliefs) and values (evaluations) in the TRA's attitude model can be traced to general theorizing about the formation and structure of social attitudes. In a paper on attitudes and motivation, Peak hypothesized that the attitude toward any object "is related to the ends which it serves, that is, to its consequences" (1955: 153) and that the attitude therefore is "some

function of (1) the judged probability that the object leads to good or bad consequences, and (2) the intensity of the affect expected from those consequences" (1955: 154). She proposed that the judged probability of a given consequence be multiplied by its expected valence and that the products be summed across all consequences to provide an estimate of the affect or evaluation associated with the object, that is, to provide an estimate of attitude (see also Carlson, 1956; Rosenberg, 1956).

Whereas Peak arrived at her attitude model by a consideration of motivation and its implication for attitude structure, Fishbein (1963) proposed a very similar model but supplied it with a concept formation and learning theory foundation (Fishbein, 1967b). His *summation model* of attitudes inspired the expectancy-value model in the TRA. In an initial test of his model, Fishbein (1963) examined the relation between beliefs about and attitudes toward African Americans. In contrast to the approach taken by Rosenberg (1956), who constructed a predetermined set [p. 442 ↓] of end states with respect to which beliefs were assessed, Fishbein elicited people's own beliefs in a free-response format. The participants were read, five times, a list of different groups of people, including "Negroes," and were asked to respond with a word they believed to be characteristic of the group in question. The ten attributes mentioned most frequently in relation to Negroes were selected for further investigation. Among these attributes were dark skin, curly hair, musical, athletic, uneducated, and hard workers.

In the second part of the study, evaluations of the ten attributes were assessed by means of five bipolar evaluative scales, including *good-bad*, *clean-dirty*, and *wise-foolish*. To measure belief strength, participants were asked to rate, for each of the ten attributes, the likelihood that blacks have the attribute in question on five probabilistic scales (e.g., *unlikely-likely*, *probably-improbably*, *false-true*). Finally, attitudes toward African Americans were assessed by means of the same five evaluative scales that had been used to measure attribute evaluations. Belief strength was multiplied by attribute evaluation and the products were summed over the ten attributes. This expectancy-value composite was found to correlate 0.80 with the direct attitude measure.

The Cognitive Foundation of Attitudes

In sum, several lines of theorizing in the 1950s and 1960s converged on the expectancy-value model of attitudes (for reviews, see Dabholkar, 1999; Feather, 1982).¹ However, the expectancy-value model of attitude in the TRA has certain features that are not necessarily shared by other expectancy-value approaches to attitude formation and structure.

Causal Effects of Beliefs on Attitudes

Perhaps most important from a theoretical perspective are differences between the TRA's attitude construct and other attitudinal approaches in the assumptions they make about the nature of the relations between beliefs and attitudes. Following Rosenberg and Hovland (1960), many investigators consider beliefs (or cognitions) and evaluations (or affect) to be two components of attitude, together with behavioral inclinations (or conation), the third component. According to this approach, attitudes can be inferred from cognitive, affective, or conative responses to the attitude object, but there is no assumption that one component causally precedes another. The tripartite model merely stipulates that there will be pressure for the three components to be evaluatively consistent with each other (Rosenberg, 1965). By way of contrast, in the TRA, beliefs that performing a behavior will lead to certain outcomes, together with the evaluations of these outcomes, are assumed to produce a favorable or unfavorable attitude toward the behavior in question. And, as we shall see below, this attitude is further assumed to have a causal effect on intentions to engage in the behavior, that is, on conation.

It is important to note, however, that not all potential outcomes of a behavior are expected to influence attitudes. According to the TRA's expectancy-value model, only beliefs that are *readily accessible* in memory determine the prevailing attitude. This limits the number of beliefs that provide the basis for an observed attitude toward a behavior and it also implies that appropriate means must be employed to identify the readily accessible beliefs. It is not sufficient, for example, to simply provide participants in a study with a list of belief statements constructed by the investigator. Many of these

statements may not represent beliefs that are readily accessible, and some accessible beliefs may be missing. Although responses to a priori set of belief statements can be used to infer underlying attitudes, it would be a mistake to assume that these responses necessarily provide information about accessible beliefs that provide the causal basis for the attitude.

[p. 443 ↓]

Subjective Norm

As noted earlier, in the TRA, Dulany's (1968) behavioral hypothesis is termed normative belief, that is, a belief that a particular referent other wants us to perform a given behavior. It is assumed that the normative beliefs regarding different social referents combine to produce an overall perceived social pressure or *subjective norm*. Drawing an analogy to the expectancy-value model of attitude toward a behavior, the prevailing subjective norm (SN) is determined by the total set of readily accessible normative beliefs concerning the expectations of important referents. Specifically, the strength of each normative belief (n_i)

) is weighted by motivation to comply (m_i)

) with referent i , and the products are aggregated across all accessible referents, as shown in Equation (21.3).

$$SN \propto \sum n_i m_i \quad (21.3)$$

We can form beliefs as to what is expected of us by being told or by inferring what important others want us to do (*injunctive norms*), or based on the observed or inferred actions of those important social referents (*descriptive norms*) (see Cialdini et al., 1990; Fishbein and Ajzen, 2010). Subjective norms are conceptually independent of attitudes toward the behavior. People can, in principle, hold favorable attitudes toward a given behavior, yet perceive social pressure not to perform it; they can hold negative attitudes

toward the behavior and favorable subjective norms; or their attitudes and subjective norms may coincide. However, in practice, personal attitudes and subjective norms are rarely, if ever, completely orthogonal to each other. This is due to the fact that many events are likely to lead to the formation of parallel behavioral and normative beliefs. Consider, for example, the publication in the popular press of the results of new medical research indicating that a low-calorie diet prolonged the life of laboratory mice by 35 percent. People exposed to this information may form the behavioral belief that eating a low-calorie diet is likely to prolong their own lives and, at the same time, also form the normative beliefs that their partners and doctors would want them to eat a low-calorie diet. As a result, attitudes toward eating a low-fat diet and subjective norms with respect to this behavior are likely to correlate with each other.

The Cognitive Foundation of Subjective Norms

The product term in the subjective norm model (see Equation [21.3]) implies that the effect of normative beliefs on subjective norms is moderated by motivation to comply. The belief that an important social referent wants us to perform a particular behavior increases perceived social pressure to do so only to the extent that we are motivated to comply with the referent in question. Similar to tests of the expectancy-value model of attitudes, tests of the subjective norm model usually involve correlating the summed products of normative belief strength times motivation to comply with a direct measure of subjective norm. Direct measures are obtained by means of items that ask participants how likely it is that important others think they should perform a behavior of interest, how likely it is that important others themselves perform or would perform the behavior, and so forth. Similar items are formulated to assess normative beliefs with respect to particular social referents. That is, participants indicate how likely they think it is that certain persons or groups of individuals (e.g., spouse, coworkers) want them to perform the behavior or themselves would perform the behavior. (The normative referents that readily come to mind in a given research population are identified by means of elicitation in a free-response format.) Finally, participants are asked to rate how motivated they are to comply with each of the

normative referents. The measures of normative belief strength are multiplied by the corresponding measures of motivation to [p. 444 ↓] comply, and the products are summed to produce a normative belief aggregate which is then correlated with the direct subjective norm measure (for illustrations, see Ajzen and Driver, 1991; Conner et al., 1998).

Empirical evidence is supportive of a correlation between an aggregate of normative beliefs on one hand and perceived social pressure or subjective norm on the other. The strength of this correlation is conveyed in a meta-analysis of research with the TPB (Armitage and Conner, 2001). Across 34 sets of data dealing with diverse kinds of behavior, the mean correlation between normative beliefs and subjective norms was 0.50. However, a number of investigators have reported that this correlation is attributable to the measure of normative belief strength and that taking motivation to comply into account does little to improve the correlation or may even lower it slightly (e.g., Ajzen and Driver, 1991; Budd et al., 1984).² One possible explanation of these findings is that people generally tend to be motivated to comply with their social referents and there is therefore relatively little meaningful variance in motivation to comply measures. Under these circumstances, multiplying normative beliefs by motivation to comply can do little to improve prediction of subjective norms.

Historical and Theoretical Context

Soon after its publication, the TRA began to stimulate a great deal of empirical research designed to predict and explain behaviors in various domains (see Sheppard et al., 1988 for an early review of this literature). The theory's broad appeal must be understood in the context of previous failures to find a strong link between verbal attitudes and actual behavior. Prior to development of the TRA, much attitude theory and research dealt with attitudes toward such general concepts as institutions, policies, racial or ethnic groups, and other broad objects. Investigators assumed that such attitudes would be predictive of any behavior toward the object of the attitude, but empirical research challenged this assumption. Attitudes toward African Americans were related neither to conformity with the judgments made by African Americans (Himelstein and Moore, 1963) nor to willingness to have a picture taken with an African

American (De Fleur and Westie, 1958; Linn, 1965); job satisfaction attitudes failed to predict job performance, absenteeism, and turnover (e.g., Bernberg, 1952; Vroom, 1964); attitudes toward labor unions failed to predict attendance at labor union meetings (Dean, 1958), and so forth. In a highly influential review of this literature, Wicker (1969) called attention to the inconsistency between verbal attitudes and overt behaviors and, like several theorists before him (e.g., Blumer, 1955; Deutscher, 1966; Festinger, 1964), questioned the utility of the attitude concept.

Compatibility

When I began my work with Martin Fishbein in the 1960s, we faced the challenge of explaining why verbal attitudes failed to predict actual behavior. In our work on the TRA, we (Ajzen, 1982; Ajzen and Fishbein, 1980; Fishbein and Ajzen, 1975) drew a distinction between two kinds of attitudes: general attitudes toward physical objects, institutions, groups, policies, and events – attitudes of the kind studied in most prior research; and attitudes toward performing particular behaviors, whether related to matters of health and safety (exercising, using contraception, getting a cancer screening, wearing a safety helmet, eating a healthy diet), race relations (hiring a member of a minority group, inviting an outgroup member to a party), politics (participating in an election, donating money to a political candidate, voting for a candidate), the environment (using public transit, recycling, conserving energy), or any other domain. We formulated the principle of correspondence or compatibility (Ajzen, 1987; Ajzen and Fishbein, 1977) to help clarify the nature of the relation [p. 445 ↓] between verbal attitudes and overt actions. According to this principle, attitudes and behavior correlate with each other to the extent that they are compatible in terms of their action, target, context, and time elements. Measures of behavior typically involve a specific action (e.g., making friends) and target (e.g., a gay person), and often also a specific context (e.g., at school) and time frame (e.g., in the next six months). By way of contrast, general attitudes (e.g., toward gays) identify only the target; they do not specify any particular action, context, or time element. We proposed that this lack of compatibility, especially in the action element, was responsible for the reported low and often nonsignificant relations between general attitudes and specific behaviors.

This is not to say, however, that general attitudes toward targets are irrelevant when it comes to the prediction of behavior. According to the principle of compatibility, general attitudes predict broad patterns or aggregates of behavior (Ajzen, 2005a; Ajzen and Fishbein, 1977). When we aggregate different behaviors directed at a given target, we generalize across actions, contexts, and time elements, thus assuring compatibility with equally broad attitudes toward the target in question. Thus, attitudes toward religion and the church, though largely unrelated to individual behaviors in this domain, were shown to correlate strongly with broad patterns of religious behavior (Fishbein and Ajzen, 1974); and attitudes toward protection of the environment predicted an aggregate of individual behaviors protective of the environment (Weigel and Newman, 1976).

The TRA succeeded where general attitudes had failed: it offered a way to use attitudes as a means to predict and explain individual behaviors. According to the theory, and in line with the principle of compatibility, individual behaviors can be predicted from attitudes toward the particular behavior of interest, and this is indeed how the attitudinal component in the TRA is defined. In addition, the theory went beyond the impact of personal attitudes by considering the role of perceived social norms, again in relation to the particular behavior of interest. Thus, the theory stipulated that the intention to perform a particular behavior is a joint function of a favorable or unfavorable attitude toward the behavior and of a subjective norm that encourages or discourages its performance.

The Theory of Planned Behavior

When we initially formulated the TRA we explicitly confined it to behaviors over which people have complete volitional control under the assumption that this category includes most behaviors of interest to social psychologists (see Ajzen and Fishbein, 1980). However, I soon came to realize that this formulation imposed too severe a limitation on a theory designed to predict and explain all manner of socially significant behavior. Many behaviors, even if in principle under volitional control, can pose serious difficulties of execution. Consider, for example, a study on physical exercise among cancer patients (Courneya et al., 2000). The investigators reported that medical complications following high-dose chemotherapy and bone marrow transplantation present significant challenges to patients' ability to adhere to a recommended exercise regimen. I decided

that, to accommodate behaviors over which people may have limited volitional control, the TRA model had to be expanded by taking degree of control over the behavior into account (Ajzen, 1985). The TPB (Ajzen, 1987, 1991b, 2005a) was designed to accomplish this goal.

Behavioral Control

Many factors, internal and external, can impair (or facilitate) performance of a given behavior: the extent to which people possess [p. 446 ↓] the requisite information, mental and physical skills and abilities, the availability of social support, emotions and compulsions, and absence or presence of external barriers and impediments (see Ajzen, 2005a: Chapter 5). People should be able to act on their intentions to the extent that they have the information, intelligence, skills, abilities, and other internal factors required to perform the behavior and to the extent that they can overcome any external obstacles that may interfere with behavioral performance. The degree of actual behavioral control is thus expected to moderate the effect of intentions on behavior. When control is uniformly high such that virtually everybody can perform the behavior if so inclined, intentions alone should be sufficient to predict behavior, but when degree of control varies among individuals, intentions and control should interact to jointly affect behavioral performance. Individuals who intend to perform the behavior and who have a high degree of control over it should be most likely to perform it.

Perceived Behavioral Control

Perhaps less self-evident than the importance of actual control, but more interesting from a psychological perspective, is the role of *perceived* behavioral control – the extent to which people believe that they can perform a given behavior if they are inclined to do so. The conceptualization of perceived behavioral control in the TPB owes much to Albert Bandura's work on self-efficacy (Bandura, 1977, 1986, 1997). In Bandura's social cognitive theory, people's beliefs about their capabilities to exercise control over events that affect their lives function as proximal determinants of human motivation and action. Bandura emphasized that self-efficacy is not a context-free global disposition

but that, instead, it “refers to beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments” (Bandura, 1997: 3). Clearly, the concept of perceived behavioral control in the TPB, though focusing on the extent to which people believe that they are capable of, or have control over, performing a given behavior, is very similar to Bandura's conception of self-efficacy.

A considerable body of research attests to the powerful effects of self-efficacy beliefs on motivation and performance (see Bandura and Locke, 2003 for a review). The strongest evidence comes from studies in which level of self-efficacy was experimentally manipulated to observe the effect on perseverance at a task and/or on task performance. Much of this research has been conducted in situations where intentions to perform the behavior of interest can be taken as given. Under these conditions, perseverance and task performance are found to increase with perceived self-efficacy. For example, Cervone and Peake (1986) had participants work on a series of intellectual problems (anagrams or cyclical graphs) that had no solution. Prior to this task, they manipulated self-efficacy beliefs by means of the anchoring and adjustment heuristic (Tversky and Kahneman, 1974). After drawing, ostensibly at random, either a relatively high number (18) or a relatively low number (4), participants were asked to indicate whether they thought they would be able to solve more, an equal number, or fewer problems than the number they had drawn, and – as a measure of self-efficacy – how many problems they thought they would be able to solve. The high anchor was found to produce a significantly higher level of perceived self-efficacy than the low anchor. The investigators then recorded how many times participants attempted to solve problems of a given type before switching to the second task. The results showed that participants in the high anchor condition persevered significantly longer on the unsolvable task than did participants in the low anchor condition, and this effect was completely mediated by measured self-efficacy.

Because the problems in this experiment had no solution, it was impossible to assess [p. 447 ↓] task performance. Other experiments have shown that manipulating self-efficacy influences not only perseverance but also actual task performance. Bandura and Adams (1977), for example, had ophidiophobics (people with a fear of snakes) undergo symbolic desensitization by visualizing threatening snake scenes while deeply relaxed. This manipulation increased the participants' sense of self-efficacy in relation to handling snakes and it also increased their subsequent ability to perform various

snake-handling tasks. Self-efficacy manipulations by means of bogus performance feedback were even found to increase pain tolerance on a compressor test (Litt, 1988) and performance on a physical endurance task (Weinberg et al., 1981).

Clearly, then, one way in which self-efficacy or perceived behavioral control can influence performance of difficult behaviors is by its effect on perseverance. The more people believe that they have the capacity to perform an intended behavior, the more likely they are to persevere and, therefore, to succeed. However, in the TPB, the role of perceived behavioral control goes beyond its effect on perseverance in at least two ways. First, the TPB is a general model designed to be applicable to any behavior, not only behaviors that individuals are motivated to perform. In fact, for most behaviors of interest to social psychologists, people vary greatly in their intentions. Some individuals intend to exercise, others do not; some intend to smoke marijuana, others do not; some intend to get a cancer screening, others have no intention to do so. In the TRA, intentions of this kind were said to be a function of attitudes and of subjective norms with respect to the behavior of interest. In the TPB, perceived behavioral control is added as a third determinant of behavioral intentions. Specifically, the more favorable people's attitudes and subjective norms, and the more they believe that they are capable of performing the behavior, the stronger should be their behavioral intentions. Conversely, people who do not believe that they are capable of performing the behavior under investigation will be unlikely to form an intention to do so.

Perceived behavioral control can thus influence behavioral performance indirectly by its effects on intentions to engage in the behavior and on perseverance in the face of difficulties encountered during execution. In addition, perceived behavioral control can potentially serve as a proxy for actual control. Recall that actual control is expected to moderate the effect of intentions on behavior. However, in most TPB applications, measures of actual control are unavailable. Indeed, with respect to many behaviors, it would be difficult to identify, let alone measure, the various internal and external factors that may facilitate or inhibit behavioral performance. It is perhaps for this reason that many investigators rely on measures of perceived behavioral control. This of course assumes that perceptions of behavioral control accurately reflect the person's actual control in the situation. To the extent that perceptions of control are veridical, they can serve as a proxy for actual control and contribute to the prediction of behavior.

The Cognitive Foundation of Perceived Behavioral Control

Like attitudes and subjective norms, perceptions of behavioral control are assumed to follow consistently from readily accessible beliefs, in this case beliefs about resources and obstacles that can facilitate or interfere with performance of a given behavior. Analogous to the expectancy-value model of attitudes, the power of each control factor to facilitate or inhibit behavioral performance is expected to contribute to perceived behavioral control in direct proportion to the person's subjective probability that the control factor is present. This model is shown in Equation (21.4), where PBC is perceived behavioral control, c

i

is the subjective probability or belief that control factor i is present, p

i

is the power of control factor i to facilitate or inhibit performance of the behavior, and [p. 448 ↓] the sum is over the total number of accessible control beliefs.

$$PBC \propto \sum c_i p_i \quad (21.4)$$

In support of this model, empirical evidence shows a strong correlation between a direct measure of perceived behavioral control and a composite of control beliefs. Direct measures of perceived behavioral control are typically obtained by asking people whether they believe that they are capable of performing the behavior of interest, whether they believe that doing so is completely under their control, and so forth. Readily accessible control factors are elicited in a free-response format. To illustrate, in a study on eating a low-fat diet (Armitage and Conner, 1999), the seven most frequently mentioned control factors dealt largely with obstacles to maintaining a low-fat diet: that doing so is time-consuming, expensive, and inconvenient; that it requires strong motivation and knowledge of the fat contents of various foods; that low-fat foods must be readily available; and that high-fat foods pose temptation.

Although investigators have frequently measured control belief strength; that is, the subjective probabilities that certain control factors will be present, in only a few studies have they also secured measures of the power of these control factors to facilitate or inhibit behavioral performance. Nevertheless, the results of empirical research provide support for the proposition that perceived behavioral control can be predicted from control beliefs. For example, in an analysis of 16 of their own studies in the health domain, Gagné and Godin (2000) found a median correlation of 0.57 between control belief composites and direct measures of perceived behavioral control, and in a meta-analysis of 18 studies on a variety of different behaviors, Armitage and Conner (2001) reported a mean correlation of 0.52.

Predicting Intentions and Behavior

To summarize briefly, according to the TPB, human action is guided by three kinds of considerations: readily accessible beliefs about the likely outcomes of the behavior and the evaluations of these outcomes (behavioral beliefs), readily accessible beliefs about the normative expectations and actions of important referents and motivation to comply with these referents (normative beliefs), and readily accessible beliefs about the presence of factors that may facilitate or impede performance of the behavior and the perceived power of these factors (control beliefs). In their respective aggregates, behavioral beliefs produce a favorable or unfavorable attitude toward the behavior; normative beliefs result in perceived social pressure or subjective norm; and control beliefs give rise to perceived behavioral control. In combination, attitude toward the behavior, subjective norm, and perception of behavioral control lead to the formation of a behavioral intention. As a general rule, the more favorable the attitude and subjective norm, and the greater the perceived control, the stronger the person's intention to perform the behavior in question. Finally, given a sufficient degree of actual control over the behavior, people are expected to carry out their intentions when the opportunity arises. Intention is thus assumed to be the immediate antecedent of behavior. However, because many behaviors pose difficulties of execution, it is useful to consider perceived behavioral control in addition to intention. Perceived control influences perseverance in the face of difficulties and, to the extent that it is veridical, it can serve as a proxy for actual control and contribute to the prediction of behavior.

Accessibility of Beliefs

I have emphasized the idea that it is the currently accessible behavioral, normative, and control beliefs that provide the cognitive foundation for attitudes, subjective norms, and perceived control, respectively. Although belief accessibility was an important feature of my theoretical framework from the outset, I became fully aware of its various [p. 449 ↓] implications only over time. In terms of explaining attitudes and behavior, the importance of belief accessibility is readily apparent. When we identify people's readily accessible beliefs we obtain a snapshot of the kinds of considerations that guide their attitudes, subjective norms, and perceptions of control and that therefore influence their intentions and actions at a given point in time. What may be less obvious is that the beliefs that are readily accessible in memory can change over time. This possibility can help explain frequently observed gaps between intentions and actions. Intentions measured at time 1 will be influenced by the beliefs that are accessible at that time. Behavior, however, is performed at a later point in time, and at that time different beliefs may have become accessible, producing different intentions. In short, intentions measured at time 1 can be expected to predict behavior at time 2 only to the extent that the same beliefs (or beliefs of equivalent valence) are readily accessible at the two points in time (Ajzen and Sexton, 1999).

Predicting Behavior

Fundamental to the TPB and other reasoned action models is the idea that behavior is guided by intentions. This idea implies, first, a strong relation between intentions and behavior, though this relation can be moderated by the degree of control over performance of the behavior. Second, it implies that changes in intentions will be followed by changes in behavior. There is ample evidence for both of these propositions.

Many studies have shown that behavioral intentions account for a considerable proportion of variance in behavior. To give just one example from our own research program, in an application of the TPB to outdoor recreation activities (Hrubec et al.,

2001), a correlation of 0.62 was observed between intentions to go hunting and self-reported hunting behavior. Meta-analyses of research in diverse behavioral domains – from physical activity, health screening, and illicit drug use to playing videogames, donating blood, and smoking cigarettes – have reported mean intention-behavior correlations ranging from 0.44 to 0.62 (e.g., Armitage and Conner, 2001; Notani, 1998; Randall and Wolff, 1994; Sheppard et al., 1988). In a meta-analysis of such meta-analyses, Sheeran (2002) reported an overall mean correlation of 0.53 between intention and behavior.

As noted previously, perceived behavioral control is expected to moderate the relation between intentions and behavior such that intentions will predict behavior better when perceived control is high rather than low. Tests of this hypothesis, when undertaken at all, typically rely on multiple regression analyses in which intentions and perceived control are entered on the first step and the product of these variables on the second step. In many of these tests the interaction term does not reach conventional levels of significance and even when it does have a significant regression coefficient, it tends to accounts for relatively little additional variance in the prediction of behavior (see Ajzen, 1991b; Armitage and Conner, 2001; Yang-Wallentin et al., 2004).

A likely explanation for such findings is the fact that there is relatively little variance in perceived behavioral control for many behaviors studied by social psychologists. Although people vary greatly in their intentions to engage in physical activity, eat a healthy diet, donate blood, recycle glass and paper, attend church, vote in an upcoming election, drink alcohol, attend class, and so forth, most believe that they can, if they so desire, engage in these kinds of activities. Under these conditions, intentions will have good predictive validity, but we cannot expect that perceived behavioral control will exert a strong moderating effect on the correlation between intentions and behavior.

With respect to some behaviors, the situation seems to be reversed such that people generally intend to perform the behavior in question but they vary greatly in their perceived control. We noted examples of this [p. 450 ↓] situation earlier in the discussion of research on self-efficacy. Participants in this research are asked to perform an intellectual or physical task, to overcome certain phobias, to tolerate pain, and so forth. It can be assumed that they try their best in these situations, that is, that they intend to perform the task to the best of their abilities. When this is the case,

behavioral achievement is found to covary with perceived self-efficacy, that is, with perceived control over the behavior. Although intentions are usually not assessed in these studies, it stands to reason that measures of intention would show relatively little variance and hence would not make much of a contribution to the prediction of behavior, either as a main effect or in interaction with perceived behavioral control.

In short, we cannot expect to find a strong interaction between intentions and perceived behavioral control when there is relatively little variance in either of these factors. Only when, in the population under investigation, people vary greatly in their intentions to perform the behavior of interest and vary in their perceptions of control over the behavior, can we expect a strong moderating effect. (For a discussion of a parallel problem in relation to the interaction of beliefs and evaluations in the expectancy-value model of attitudes, see Ajzen and Fishbein, 2008.)

Causal Effect of Intentions on Behavior

In reasoned action models, such as the TPB, intentions are assumed to be causal antecedents of corresponding behavior. The correlational nature of most empirical evidence shows that intentions can indeed be used to predict behavior, but such evidence is not definitive proof of their causal impact. There is growing evidence, however, for a causal effect of intentions on actions coming mainly from intervention studies. In a meta-analysis of 47 studies in which an intervention was shown to have had a significant effect on intentions (Webb and Sheeran, 2006), this effect was also shown to promote a change in actual behavior. On average, the interventions reviewed produced medium to large changes in intentions (mean $d = 0.66$), followed by small to medium changes in behavior (mean $d = 0.36$).

Predicting and Explaining Intentions

There is also ample evidence to show that, consistent with the TPB, intentions can be predicted from attitudes, subjective norms, and perceptions of behavioral control. Two examples will serve to illustrate successful applications of the theory. In the study by Hrubes et al. (2001) mentioned earlier, the multiple correlation for the prediction of

hunting intentions was 0.92, showing that attitudes, subjective norms, and perceived control accounted for 86 percent of the variance in intentions. Each of the three antecedents of intentions made a significant contribution to the prediction although attitudes were found to be most important ($\beta = 0.58$), followed by subjective norms ($\beta = 0.37$) and perceived control ($\beta = 0.07$). A different pattern of influence was observed in a study on leisure time physical activity among individuals with spinal cord injury (Latimer and Martin Ginis, 2005). The multiple correlation for the prediction of intentions was 0.78, indicating that attitudes, subjective norms, and perceptions of behavioral control accounted for 61 percent of the variance in intentions. Again, the regression coefficients were statistically significant for all three predictors. However, perhaps not surprisingly given the difficulties individuals with spinal cord injuries are likely to face, perceived control made a larger independent contribution to the prediction of exercise intentions ($\beta = 0.46$) than did either attitudes ($\beta = 0.29$) or subjective norms ($\beta = 0.27$).

It is beyond the scope of this chapter to review the large body of research attesting to the proposition that intentions can be predicted from attitudes toward the behavior, subjective norms, and perceived behavioral control. In meta-analytic syntheses covering [p. 451 ↓] varied behaviors (Armitage and Conner, 2001; Cheung and Chan, 2000; Notani, 1998; Rivas and Sheeran, 2003; Schulze and Wittmann, 2003), the mean multiple correlations for the prediction of intentions ranged from 0.59 to 0.66. Meta-analyses in specific behavioral domains show similar results. In two meta-analytic reviews of research on condom use, the mean multiple correlations were found to be 0.71 (Albarracín et al., 2001) and 0.65 (Sheeran and Taylor, 1999), and in two meta-analyses of research on physical activity, the mean multiple correlations were 0.55 (Downs and Hausenblas, 2005) and 0.67 (Hagger et al., 2002). An extensive review of this literature can be found in Fishbein and Ajzen (2010, chapter 6).

Reasoned Action

The TPB emphasizes the controlled aspects of human information processing and decision making. Its concern is primarily with behaviors that are goal-directed and steered by conscious self-regulatory processes. According to the TPB, intentions and behavior are guided by expected consequences of performing the behavior, by

perceived normative pressures, and by anticipated difficulties. This focus has often been misinterpreted to mean that the theory posits an impassionate, rational actor who reviews all available information in an unbiased fashion to arrive at a behavioral decision. In reality, the theory draws a much more complex and nuanced picture.

First, there is no assumption in the TPB that behavioral, normative, and control beliefs are formed in a rational, unbiased fashion or that they accurately represent reality. Beliefs reflect the information people have in relation to the performance of a given behavior, but this information is often inaccurate and incomplete; it may rest on faulty or irrational premises, be biased by self-serving motives, or otherwise fail to reflect reality. Clearly, this is a far cry from a rational actor. However, no matter how people arrive at their behavioral, normative, and control beliefs, their attitudes toward the behavior, their subjective norms, and their perceptions of behavioral control follow automatically and consistently from their beliefs. It is only in this sense that behavior is said to be reasoned or planned. Even if inaccurate, biased, or otherwise irrational, our beliefs produce attitudes, intentions, and behaviors consistent with these beliefs (see, e.g., Geraerts et al., 2008).

Second, there is no assumption in the TPB that people carefully and systematically review all their beliefs every time they are about to perform a behavior. On the contrary, the theory recognizes that most behaviors in everyday life are performed without much cognitive effort. Consistent with contemporary theorizing in social psychology (see Carver and Scheier, 1998; Chaiken and Trope, 1999; Petty and Cacioppo, 1986), it is assumed that the amount of information processing people engage in prior to performing a behavior varies along a continuum, from shallow to deep (Ajzen and Sexton, 1999). In-depth processing is reserved for important decisions and behaviors in novel situations that demand careful consideration of the behavior's likely consequences, the normative expectations of significant others, and the obstacles that may be encountered. When it comes to routine, everyday behaviors like eating breakfast, taking one's vitamin supplements, going to work, watching the news on TV, and so forth, no careful deliberation is required or postulated. Attitudes, subjective norms, and perceptions of control as well as intentions in relation to these kinds of behaviors are assumed to guide behavior implicitly without cognitive effort and often below conscious awareness (see Ajzen and Fishbein, 2000 for a discussion of these issues).

Habituation and Automaticity in Social Behavior

Notwithstanding the above qualifications, the reasoned action approach represented in the [p. 452 ↓] TPB stands in contrast to recent trends in social psychology that view much of human social behavior as habitual, automatic, and driven by nonconscious goal pursuit (Bargh, 1990; Bargh and Barndollar, 1996; Bargh et al., 2001; Hassin et al., 2009; Kruglanski et al., 2002; Ouellette and Wood, 1998). In this section, I briefly consider the issues raised and how they can be reconciled with a reasoned action perspective.

Habituation

With repeated performance, behavior is said to habituate such that it comes under direct control of stimulus cues, bypassing intentions as a determinant of behavior. This argument implies that behavioral intentions lose their predictive validity once a strong habit has been established (e.g., Aarts et al., 1998; Neal et al., 2006; Ouellette and Wood, 1998). Empirical findings lend little support to this hypothesis. In a meta-analysis of 15 data sets, Ouellette and Wood (1998) classified each set as dealing with a behavior that can be performed frequently and hence can habituate (e.g., seatbelt use, coffee drinking, class attendance) or infrequently and not likely to habituate (e.g., flu shots, blood donation, nuclear protest). Contrary to the habit hypothesis, prediction of behavior from intentions was found to be quite accurate for both types of behavior (mean $r = 0.59$ and $r = 0.67$ for high- and low-opportunity behaviors, respectively; difference not significant). The same conclusion comes from a more extensive meta-analysis based on 51 data sets (Sheeran and Sutton, unpublished data). For behaviors that could be performed infrequently (once or twice a year) the intention-behavior correlation was 0.51, and it was 0.53 for high-opportunity behaviors that could be performed daily or at least once a week. This meta-analysis also compared behaviors typically performed in the same context and thus amenable to habit formation to behaviors performed in variable contexts. Again, there was little difference in the

predictive validity of intentions. If anything, the pattern of results was contrary to what would be predicted by the habit hypothesis. The mean intention-behavior correlation was 0.40 with respect to behaviors performed in unstable contexts (where intentions should be most relevant) compared with a mean intention-behavior correlation of 0.56 for behaviors performed in stable contexts. Nor were the results more supportive of the habit hypothesis in a direct test with primary data (Ouellette and Wood, 1998; see Ajzen, 2002 for a discussion of these issues).

In sum, although behavior can become routine with repeated performance, no longer requiring much if any conscious deliberation, there is no evidence to suggest that intentions become irrelevant when behavior is routine. On the contrary, empirical evidence demonstrates that intentions predict routine as well as relatively novel behaviors. Moreover, this conclusion is not necessarily inconsistent with a habit perspective. “Within current theorizing, habits are automated response *dispositions* that are cued by aspects of the performance context” (Neal et al., 2006: 198; emphasis added). It may thus be argued that it is not the routinized behavior itself that is automatically initiated by the supporting context but a disposition to perform the behavior, such as an implicit intention. Consistent with this idea, the TPB assumes that, in the case of routine behaviors, implicit intentions are activated automatically and are then available to guide performance of the behavior.

Automaticity

Priming research has demonstrated that a large array of psychological concepts and processes can be activated automatically, below conscious awareness (see Bargh, 2006). Initial studies showed that activation of such knowledge structures as trait concepts (kindness, hostility) or ethnic stereotypes can influence encoding, comprehension, and judgments of ambiguous social behavior (e.g., Higgins et al., 1977; Srull and Wyer, 1979). In recent years, research has turned to nonconscious goal pursuit, demonstrating that desired [p. 453 ↓] outcomes can be primed subconsciously and influence pursuit of the activated goal without the person's awareness (Hassin et al., 2009; Kruglanski et al., 2002). Thus, for example, primed activation of an achievement goal was found to improve level of performance on a word-search task (Bargh et al., 2001).

More relevant for present purposes, however, is the proposition that automatic activation of knowledge structures or goals can influence not only judgments or achievements but can also have a direct effect on behavior. Bargh et al. (1996), for example, showed that participants primed with the stereotype of the elderly walked more slowly down the hallway when leaving the experiment than did control participants, and when primed with the concept of rudeness, they interrupted the experimenter more frequently and quickly than when primed with the concept of politeness. Similarly, Aarts and Dijksterhuis (2003) found that when primed with the concept of silence by exposure to a picture of a library, participants spoke more softly, and following exposure to a picture of an exclusive restaurant, they were more likely to remove crumbs after eating a biscuit. These effects have usually been attributed to automatic enactment of a response made ready accessible by priming a certain construct (ideo-motor expression).

Traditionally, goal-directed behavior has been conceptualized as a controlled process that involves some measure of conscious deliberation and awareness, a view inherent in the TPB and other reasoned action models (e.g., Bandura, 1986, 1997; Deci and Ryan, 1985; Locke and Latham, 1990; Triandis, 1977). Although consistent with our intuitive sense that the pursuit of nontrivial goals is a controlled, conscious process, this perspective has, in the past 20 years, given way to theorizing that denies the importance of consciousness as a causal agent (Wegner, 2002; Wegner and Wheatley, 1999) and views much human social behavior as driven by implicit attitudes (Greenwald and Banaji, 1995) and other unconscious or subconscious mental processes (Aarts and Dijksterhuis, 2000; Bargh, 1989, 1996; Bargh and Chartrand, 1999; Brandstätter et al., 2001; Uhlmann and Swanson, 2004).

In one sense, the importance of nonconscious processes is undeniable. Our ability to exercise conscious, intentional control is constrained by limited information processing capacity such that most moment-to-moment mental processes must occur below conscious awareness (Bargh and Chartrand, 1999). The question of how much of our day-to-day behaviors is subject to automatic versus controlled processes is complicated by the fact that behavior involves a complex sequence of events. Many attributes of behavioral performance are outside conscious awareness. Thus, we don't pay much attention to how we move our legs and arms as we walk or how we produce sentences as we speak, nor do we ordinarily consciously monitor our facial expressions, tone

of voice, or body posture. Even more complex behaviors can become automatic with sufficient practice. When we learn to drive, for example, we initially pay close attention to various aspects of this behavior, but once we have become skilled at the task, we can perform it more or less automatically as evidenced by the fact that we can at the same time engage in conversation or other activities that occupy our cognitive resources.

The research reviewed above has demonstrated that priming of constructs, knowledge structures, and goals can initiate these kinds of automatic processes and routine action sequences, although questions can be raised as to whether important decisions, such as buying a car, are ever completely automatic. More importantly, it should be noted that the observed automaticity in behavior is also consistent with a reasoned action perspective if we assume that attitudes and intentions with respect to common behaviors can become implicit and exert their influence below conscious awareness. Recent empirical research (e.g., Cesario et al., 2006; see also Förster et al., 2005) has provided support for this idea by showing that the priming of a category activates implicit preparatory responses, such as an implicit attitude toward [p. 454 ↓] the primed category, and that these implicit responses determine the effects of the prime on behavior. For example, Cesario et al. (2006, Study 2) assessed implicit attitudes toward elderly and youth and, in a second session, subliminally presented either elderly picture primes or youth picture primes (or no primes in the control condition). Following this manipulation, the participants' speed of walking was recorded. As in previous research, priming the elderly construct slowed walking speed whereas priming the youth construct increased it. Of greater interest, however, was the role of implicit attitudes. The more positive participants' implicit attitudes toward the elderly, the slower they walked; and the more negative their implicit attitudes toward the elderly, the faster they walked. Comparable results were obtained with respect to activation of implicit attitudes toward youth. Speed of walking was therefore not a simple automatic response directly produced by priming of the elderly or youth category. Instead, in a proximal sense, it was the result of implicit preparatory responses, that is, of implicit attitudes toward the elderly or toward youth that were activated by priming these categories.

Of course, theorists interested in automaticity do not deny the importance of controlled processes in social attitudes and behavior (Devine and Monteith, 1999; Wegner and Bargh, 1998). Most endorse a dual-mode processing perspective that has room for

automatic as well as controlled processes (Chaiken and Trope, 1999), but in recent years the pendulum may have swung too much in the direction of automaticity. If history is a guide, the pendulum is eventually bound to swing in the opposite direction and perhaps the time has come for social psychologists to rediscover reasoned action.

Applications of the Theory

Looking back over the past 30 years, I am very gratified to see that the TPB has proven to be a useful framework for understanding, predicting, and changing human social behavior. Judging by the sheer number of investigations it has stimulated, the TPB is perhaps the most popular of the reasoned action models (for a list of publications, see bibliography in Ajzen, 2005b). Its application in varied domains has allowed investigators to identify important psychological determinants of socially significant behaviors. Armed with the conceptual framework and methodologies provided by the TPB, investigators have collected information about the behavioral, normative, and control-related determinants of many different behaviors, from exercising, eating a healthy diet, donating blood, and using illicit drugs to conserving energy, using public transportation, and practicing safer sex. Such knowledge can, of course, also provide the basis for effective interventions designed to modify social behavior in a desirable direction. Although the number of actual intervention studies to date is relatively small, especially in comparison with the large number of prediction studies, the theory has demonstrated its utility as a basis for designing and evaluating the effectiveness of interventions of various kinds, including interventions to discourage car use (Bamberg and Schmidt, 2001), limit infant sugar intake (Beale and Manstead, 1991), promote effective job search behaviors (Van Ryn and Vinokur, 1992), and encourage testicular self-examination (Brubaker and Fowler, 1990) and condom use (Fishbein et al., 1997) (see Ajzen, *in press*, for a review). I am hopeful that, in this fashion, the theory will continue to make a valuable contribution to the solutions of critical social problems.

Notes

1 A popular approach to the analysis of decisions under uncertainty, the subjective expected utility model of behavioral decision theory (Coombs and Beardslee, 1954; Edwards, 1954), was developed at about the same time. According to this model, the [p. 455 ↓] expected utility of a choice alternative is a function of the subjective probability that the alternative possesses certain attributes, multiplied by the subjective values or utilities of those attributes. It is assumed that a subjective expected utility is produced for each choice alternative and that decision makers choose the alternative with the highest SEU. For a comparison of the expectancy-value and SEU models, see Ajzen (1996).

2 Some investigators (e.g., Fekadu and Kraft, 2002; Rimal and Real, 2003) have assessed identification with the social referent, instead of motivation to comply, and have examined the moderating effect of identification. These investigations have shown that identification with social referents also does little to moderate the effect of normative beliefs.

References

Aarts, H. and Dijksterhuis, A. (2000) Habits as knowledge structures: Automaticity in goal-directed behavior . *Journal of Personality and Social Psychology* , 78, 53–63.
<http://dx.doi.org/10.1037/0022-3514.78.1.53>

Aarts, H. and Dijksterhuis, A. (2003) The silence of the library: Environment, situational norm, and social behavior . *Journal of Personality and Social Psychology* , 84, 18–28.
<http://dx.doi.org/10.1037/0022-3514.84.1.18>

Aarts, H., Verplanken, B. and van Knippenberg, A. (1998) Predicting behavior from actions in the past: Repeated decision making or a matter of habit? *Journal of Applied Social Psychology* , 28, 1355–1374. <http://dx.doi.org/10.1111/j.1559-1816.1998.tb01681.x>

Ajzen, I. (1982) On behaving in accordance with one's attitudes . In M.P. Zanna, ed. , E.T. Higgins, ed. and C.P. Herman (eds), *Consistency in Social Behavior: The Ontario Symposium* , 2, 3–15. Hillsdale, NJ: Erlbaum.

Ajzen, I. (1985) From intentions to actions: A theory of planned behavior . In J. Kuhl, ed. and J. Beckman (eds), Action-control: From Cognition to Behavior , pp. 11–39. Heidelberg: Springer. http://dx.doi.org/10.1007/978-3-642-69746-3_2

Ajzen, I. (1987) Attitudes, traits, and actions: Dispositional prediction of behavior in personality and social psychology . In L. Berkowitz (ed.) Advances in Experimental Social Psychology, Vol . 20, pp. 1–63. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960411-6>

Ajzen, I. (1991a) Benefits of leisure: A social psychological perspective . In B.L. Driver, ed. and P.J. Brown (eds), Benefits of Leisure , pp. 411–417. State College: Venture Publishing.

Ajzen, I. (1991b) The theory of planned behavior . Organizational Behavior and Human Decision Processes , 50, 179–211. <http://dx.doi.org/10.1016/0749-5978%2891%2990020-T>

Ajzen, I. (1996) The social psychology of decision making . In E.T. Higgins, ed. and A.W. Kruglanski (eds), Social Psychology: Handbook of Basic Principles , pp. 297–325. New York: Guilford Press.

Ajzen, I. (2002) Residual effects of past on later behavior: Habituation and reasoned action perspectives . Personality and Social Psychology Review , 6, 107–122. http://dx.doi.org/10.1207/S15327957PSPR0602_02

Ajzen, I. (2005a) Attitudes, Personality, and Behavior , 2nd Edition. Maidenhead: Open University Press.

Ajzen, I. (2005b) The Theory of Planned Behavior . Available at: <http://www.people.umass.edu/aizen/tpbrefs.html>, accessed 3 March 2005.

Ajzen, I. (in press) Behavioral interventions: Design and evaluation guided by the theory of planned behavior. In M.M. Mark, ed. , S.I. Donaldson, ed. and B. Campbell (eds), Social Psychology and Program/Policy Evaluation . New York: Guilford Press.

Ajzen, I. and Driver, B.L. (1991) Prediction of leisure participation from behavioral, normative, and control beliefs: An application of the theory of planned behavior . *Leisure Sciences* , 13, 185–204. <http://dx.doi.org/10.1080/01490409109513137>

Ajzen, I. and Fishbein, M. (1977) Attitude-behavior relations: A theoretical analysis and review of empirical research . *Psychological Bulletin* , 84, 888–918. <http://dx.doi.org/10.1037/0033-2909.84.5.888>

Ajzen, I. and Fishbein, M. (1980) *Understanding Attitudes and Predicting Social Behavior* . Englewood-Cliffs, NJ: Prentice-Hall.

Ajzen, I. and Fishbein, M. (2000) Attitudes and the attitude-behavior relation: Reasoned and automatic processes . In W. Stroebe, ed. and M. Hewstone (eds), *European Review of Social Psychology* , 11, 1–33. Chichester: Wiley. <http://dx.doi.org/10.1080/14792779943000116>

Ajzen, I. and Fishbein, M. (2008) Scaling and testing multiplicative combinations in the expectancy-value model of attitudes . *Journal of Applied Social Psychology* , 33, 2222–2247. <http://dx.doi.org/10.1111/j.1559-1816.2008.00389.x>

Ajzen, I. and Sexton, J. (1999) Depth of processing, belief congruence, and attitude-behavior correspondence . In S. Chaiken, ed. and Y. Trope (eds), *Dual-process Theories in Social Psychology* , pp. 117–138. New York: Guilford.

Albarracín, D., Johnson, B.T., Fishbein, M. and Muellerleile, P.A. (2001) Theories of reasoned action and planned behavior as models of condom use: A meta-analysis . *Psychological Bulletin* , 127, 142–161.

Armitage, C.J. and Conner, M. (1999) Distinguishing perceptions of control from self-efficacy: Predicting consumption of a low-fat diet using the theory of planned behavior . *Journal of Applied Social Psychology* , 29, 72–90. <http://dx.doi.org/10.1111/j.1559-1816.1999.tb01375.x>

Armitage, C.J. and Conner, M. (2001) Efficacy of the theory of planned behavior: A meta-analytic review . *British Journal of Social Psychology* , 40, 471–499. <http://dx.doi.org/10.1348/014466601164939>

Armitage, C.J., Conner, M., Loach, J. and Willetts, D. (1999) Different perceptions of control: Applying an extended theory of planned behavior to legal and illegal drug use . Basic and Applied Social Psychology , 21, 301–316.

Bamberg, S. and Schmidt, P. (2001) Theory-driven subgroup-specific evaluation of an intervention to reduce private car use . Journal of Applied Social Psychology , 31, 1300–1329. <http://dx.doi.org/10.1111/j.1559-1816.2001.tb02675.x>

Bandura, A. (1977) Self-efficacy: Toward a unifying theory of behavioral change . Psychological Review , 84, 191–215. <http://dx.doi.org/10.1037/0033-295X.84.2.191>

Bandura, A. (1986) Social Foundations of Thought and Action: A Social Cognitive Theory . Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997) Self-efficacy: The Exercise of Control . New York: Freeman.

Bandura, A. and Adams, N.E. (1977) Analysis of self-efficacy theory of behavioral change . Cognitive Therapy and Research , 1, 287–310. <http://dx.doi.org/10.1007/BF01663995>

Bandura, A. and Locke, E.A. (2003) Negative self-efficacy and goal effects revisited . Journal of Applied Psychology , 88, 87–99. <http://dx.doi.org/10.1037/0021-9010.88.1.87>

Bargh, J.A. (1989) Conditional automaticity: Varieties of automatic influence in social perception and cognition . In J.S. Uleman, ed. and J.A. Bargh (eds), Unintended Thought , pp. 3–51. New York: Guilford Press.

Bargh, J.A. (1990) Auto-motives: Preconscious determinants of social interaction . In E.T. Higgins, ed. and R.M. Sorrentino (eds), Handbook of Motivation and Cognition: Foundations of Social Behavior , pp. 93–130. New York: Guilford Press.

Bargh, J.A. (1996) Automaticity in social psychology . In E.T. Higgins, ed. and A.W. Kruglanski (eds), Social Psychology: Handbook of Basic Principles , pp. 169–183. New York: Guilford Press.

Bargh, J.A. (2006) Agenda 2006: What have we been priming all these years? On the development, mechanisms, and ecology of nonconscious social behavior . European Journal of Social Psychology , 36, 147–168. <http://dx.doi.org/10.1002/ejsp.336>

Bargh, J.A. and Barndollar, K. (1996) Automaticity in action: The unconscious as repository of chronic goals and motives . In P.M. Gollwitzer, ed. and J.A. Bargh (eds), The Psychology of Action: Linking Cognition and Motivation to Behavior , pp. 457–481. New York: Guilford Press.

Bargh, J.A. and Chartrand, T.L. (1999) The unbearable automaticity of being . American Psychologist , 54, 462–479. <http://dx.doi.org/10.1037/0003-066X.54.7.462>

Bargh, J.A., Chen, M. and Burrows, L. (1996) Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action . Journal of Personality and Social Psychology , 71, 230–244. <http://dx.doi.org/10.1037/0022-3514.71.2.230>

Bargh, J.A., Gollwitzer, P.M., Lee-Chai, A., Barndollar, K. and Troetschel, R. (2001) The automated will: Nonconscious activation and pursuit of behavioral goals . Journal of Personality and Social Psychology , 81, 1014–1027. <http://dx.doi.org/10.1037/0022-3514.81.6.1014>

Beale, D.A. and Manstead, A.S.R. (1991) Predicting mothers' intentions to limit frequency of infants' sugar intake: Testing the theory of planned behavior . Journal of Applied Social Psychology , 21, 409–431. <http://dx.doi.org/10.1111/j.1559-1816.1991.tb00528.x>

Bernberg, R.E. (1952) Socio-psychological factors in industrial morale: I. The prediction of specific indicators . Journal of Social Psychology , 36, 73–82. <http://dx.doi.org/10.1080/00224545.1952.9919090>

Blumer, H. (1955) Attitudes and the social act . Social Problems , 3, 59–65. <http://dx.doi.org/10.2307/798705>

Brandstätter, V., Lengfelder, A. and Gollwitzer, P.M. (2001) Implementation intentions and efficient action initiation . Journal of Personality and Social Psychology , 81, 946–960.

Brewer, W.F. (1974) There is no convincing evidence for operant or classical conditioning in adult humans . In W.B. Weimer, ed. and D.S. Palermo (eds), *Cognition and the Symbolic Processes* . Hillsdale, NJ: Lawrence Erlbaum Associates.

Brubaker, R.G. and Fowler, C. (1990) Encouraging college males to perform testicular self-examination: Evaluation of a persuasive message based on the revised theory of reasoned action . *Journal of Applied Social Psychology* , 20, 1411–1422. <http://dx.doi.org/10.1111/j.1559-1816.1990.tb01481.x>

Budd, R.J., North, D. and Spencer, C. (1984) Understanding seat-belt use: A test of Bentler and Speckart's extension of the 'theory of reasoned action' . *European Journal of Social Psychology* , 14, 69–78. <http://dx.doi.org/10.1002/ejsp.2420140106>

Carlson, E.R. (1956) Attitude change through modification of attitude structure . *Journal of Abnormal and Social Psychology* , 52, 256–261. <http://dx.doi.org/10.1037/h0048141>

Carver, C.S. and Scheier, M.F. (1998) *On the Self-regulation of Behavior* . Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139174794>

Cervone, D. and Peake, P.K. (1986) Anchoring, efficacy, and action: The influence of judgmental heuristics on self-efficacy judgments and behavior . *Journal of Personality and Social Psychology* , 50, 492–501. <http://dx.doi.org/10.1037/0022-3514.50.3.492>

Cesario, J., Plaks, J.E. and Higgins, E.T. (2006) Automatic social behavior as motivated preparation to interact . *Journal of Personality and Social Psychology* , 90, 893–910. <http://dx.doi.org/10.1037/0022-3514.90.6.893>

Chaiken, S., ed. and Trope, Y. (eds) (1999) *Dual-process Theories in Social Psychology* . New York: Guilford Press.

Cheung, S.-F. and Chan, D.K.-S. (2000) The role of perceived behavioral control in predicting human behavior: A meta-analytic review of studies on the theory of planned behavior . Unpublished manuscript, Chinese University of Hong Kong.

Cialdini, R.B., Reno, R.R. and Kallgren, C.A. (1990) A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public

places . Journal of Personality and Social Psychology , 58, 1015–1026. <http://dx.doi.org/10.1037/0022-3514.58.6.1015>

Conner, M., Sherlock, K. and Orbell, S. (1998) Psychosocial determinants of ecstasy use in young people in the UK . British Journal of Health Psychology , 3, 295–317. <http://dx.doi.org/10.1111/j.2044-8287.1998.tb00576.x>

Coombs, C.H. and Beardslee, D. (1954) On decision-making under uncertainty . In R.M. Thrall, ed. , C.H. Coombs, ed. and R.L. Davis (eds), Decision Processes , pp. 255–285. New York: Wiley.

Courneya, K.S., Keats, M.R. and Turner, R. (2000) Social cognitive determinants of hospital-based exercise in cancer patients following high-dose chemotherapy and bone marrow transplantation . International Journal of Behavioral Medicine , 7, 189–203. http://dx.doi.org/10.1207/S15327558IJBM0703_01

Dabholkar, P.A. (1999) Expectancy-value models . In P.E. Earl, ed. and S. Kemp (eds), The Elgar Companion to Consumer Research and Economic Psychology , pp. 200–208. Cheltenham: Edward Elgar.

Davis, F.D., Bagozzi, R.P. and Warshaw, P.R. (1989) User acceptance of computer technology: A comparison of two theoretical models . Management Science , 35, 982–1003. <http://dx.doi.org/10.1287/mnsc.35.8.982>

De Fleur, M.L. and Westie, F.R. (1958) Verbal attitudes and overt acts: An experiment on the salience of attitudes . American Sociological Review , 23, 667–673. <http://dx.doi.org/10.2307/2089055>

Dean, L.R. (1958) Interaction, reported and observed: The case of one local union . Human Organization , 17, 36–44.

Deci, E.L. and Ryan, R.M. (1985) The general causality orientations scale: Self-determination in personality . Journal of Research in Personality , 19, 109–134. <http://dx.doi.org/10.1016/0092-6566%2885%2990023-6>

Deutscher, I. (1966) Words and deeds: Social science and social policy . *Social Problems* , 13, 235–254. <http://dx.doi.org/10.2307/799252>

Devine, P.G. and Monteith, M.J. (1999) Automaticity and control in stereotyping . In S. Chaiken, ed. and Y. Trope (eds), *Dual-process Theories in Social Psychology* , pp. 339–360. New York: Guilford Press.

Downs, D.S. and Hausenblas, H.A. (2005) The theories of reasoned action and planned behavior applied to exercise: A meta-analytic update . *Journal of Physical Activity and Health* , 2, 76–97.

Dulany, D.E. (1962) The place of hypotheses and intentions: An analysis of verbal control in verbal conditioning . In C. Eriksen (ed.) *Behavior and Awareness* , pp. 102–129. Durham, NC: Duke University Press.

Dulany, D.E. (1968) Awareness, rules, and propositional control: A confrontation with S-R behavior theory . In D. Hornton, ed. and T. Dixon (eds), *Verbal Behavior and S-R Behavior Theory* , pp. 340–387. Englewood Cliffs, NJ: Prentice-Hall.

Edwards, W. (1954) The theory of decision making . *Psychological Bulletin* , 51, 380–417. <http://dx.doi.org/10.1037/h0053870>

Feather, N.T. (ed.) (1982) *Expectations and Actions: Expectancy-value Models in Psychology* . Hillsdale, NJ: Erlbaum.

Fekadu, Z. and Kraft, P. (2002) Expanding the theory of planned behaviour: The role of social norms and group identification . *Journal of Health Psychology* , 7, 33–43. <http://dx.doi.org/10.1177/1359105302007001650>

Festinger, L. (1964) Behavioral support for opinion change . *The Public Opinion Quarterly* , 28, 404–417. <http://dx.doi.org/10.1086/267263>

Fishbein, M. (1963) An investigation of the relationships between beliefs about an object and the attitude toward that object . *Human Relations* , 16, 233–240. <http://dx.doi.org/10.1177/001872676301600302>

Fishbein, M. (1967a) Attitude and the prediction of behavior . In M. Fishbein (ed.) Readings in Attitude Theory and Measurement , pp. 477–492. New York: Wiley.

Fishbein, M. (1967b) A behavior theory approach to the relations between beliefs about an object and the attitude toward the object . In M. Fishbein (ed.) Readings in Attitude Theory and Measurement , pp. 389–400. New York: Wiley.

Fishbein, M. and Ajzen, I. (1974) Attitudes towards objects as predictors of single and multiple behavioral criteria . Psychological Review , 81, 59–74. <http://dx.doi.org/10.1037/h0035872>

Fishbein, M. and Ajzen, I. (1975) Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research . Reading, MA: Addison-Wesley.

Fishbein, M. and Ajzen, I. (2010) Prediction and Change of Behavior: The Reasoned Action Approach . New York: Psychology Press.

Fishbein, M., Guenther-Grey, C., Johnson, W., Wolitski, R.J., McAlister, A., Rietmeijer, C.A. and O'Reilly, K (1997) Using a theory-based community intervention to reduce AIDS risk behaviors: The CDC's AIDS community demonstration projects . In M.E. Goldberg, ed. , M. Fishbein, ed. and S. Middlestadt (eds), Social Marketing: Theoretical and Practical Perspectives , pp. 123–146. Mahwah, NJ: Lawrence Erlbaum Associates.

Fisher, J.D. and Fisher, W.A. (1992) Changing AIDS-risk behavior . Psychological Bulletin , 111, 455–474. <http://dx.doi.org/10.1037/0033-2909.111.3.455>

Förster, J., Liberman, N. and Higgins, E.T. (2005) Accessibility from active and fulfilled goals . Journal of Experimental Social Psychology , 41, 220–239.

Gagné, C. and Godin, G. (2000) The theory of planned behavior: Some measurement issues concerning belief-based variables . Journal of Applied Social Psychology , 30, 2173–2193.

Geraerts, E., Bernstein, D.M., Merckelbach, H., Linders, C., Raymaeckers, L. and Loftus, E.F. (2008) Lasting false beliefs and their behavioral

consequences . Psychological Science , 19, 749–753. <http://dx.doi.org/10.1111/j.1467-9280.2008.02151.x>

Greenwald, A.G. and Banaji, M.R. (1995) Implicit social cognition: Attitudes, self-esteem, and stereotypes . Psychological Review , 102, 4–27. <http://dx.doi.org/10.1037/0033-295X.102.1.4>

Hagger, M.S., Chatzisarantis, N.L.D. and Biddle, S.J.H. (2002) A meta-analytic review of the theories of reasoned action and planned behavior in physical activity: Predictive validity and the contribution of additional variables . Journal of Sport and Exercise Psychology , 24, 3–32.

Hassin, R.R., Aarts, H., Eitam, B., Custers, R. and Kleiman, T. (2009) Non-conscious goal pursuit and the effortful control of behavior . In E. Morsella, ed. , P.M. Gollwitzer, ed. and J.A. Bargh (eds), Oxford Handbook of Human Action , pp. 549–568. New York: Oxford University Press.

Higgins, E.T., Rholes, W.S. and Jones, C.R. (1977) Category accessibility and impression formation . Journal of Experimental Social Psychology , 13, 141–154. <http://dx.doi.org/10.1016/S0022-1031%2877%2980007-3>

Himelstein, P. and Moore, J. (1963) Racial attitudes and the action of Negro and white background figures as factors in petition-signing . Journal of Social Psychology , 61, 267–272. <http://dx.doi.org/10.1080/00224545.1963.9919484>

Hrubes, D., Ajzen, I. and Daigle, J. (2001) Predicting hunting intentions and behavior: An application of the theory of planned behavior . Leisure Sciences , 23, 165–178. <http://dx.doi.org/10.1080/014904001316896855>

Kruglanski, A.W., Shah, J.Y., Fisbach, A., Friedman, R.S., Chun, W.Y. and Sleeth-Keppler, D. (2002) A theory of goal systems . In M.P. Zanna (ed.), Advances in Experimental Social Psychology , pp. 331–378. San Diego: Academic Press.

Latimer, A.E. and Martin Ginis, K.A. (2005) The theory of planned behavior in prediction of leisure time physical activity among individuals with spinal cord injury . Rehabilitation Psychology , 50, 389–396. <http://dx.doi.org/10.1037/0090-5550.50.4.389>

Linn, L.S. (1965) Verbal attitudes and overt behavior: A study of racial discrimination . Social Forces , 43, 353–364. <http://dx.doi.org/10.2307/2574765>

Litt, M.D. (1988) Self-efficacy and perceived control: Cognitive mediators of pain tolerance . Journal of Personality and Social Psychology , 54, 149–160. <http://dx.doi.org/10.1037/0022-3514.54.1.149>

Locke, E.A. and Latham, G.P. (1990) A Theory of Goal Setting and Task Performance . Englewood Cliffs, NJ: Prentice-Hall.

Locke, E.A. and Latham, G.P. (1994) Goal setting theory . In H.F.J. O'Neil, ed. and M. Drillings (eds), Motivation: Theory and Research , pp. 13–29. Hillsdale, NJ: Lawrence Erlbaum Associates.

Neal, D.T., Wood, W. and Quinn, J.M. (2006) Habits – a repeat performance . Current Directions in Psychological Science , 15, 198–202. <http://dx.doi.org/10.1111/j.1467-8721.2006.00435.x>

Notani, A.S. (1998) Moderators of perceived behavioral control's predictiveness in the theory of planned behavior: A meta-analysis . Journal of Consumer Psychology , 7, 247–271. http://dx.doi.org/10.1207/s15327663jcp0703_02

Ouellette, J.A. and Wood, W. (1998) Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior . Psychological Bulletin , 124, 54–74. <http://dx.doi.org/10.1037/0033-2909.124.1.54>

Peak, H. (1955) Attitude and motivation . In M.R. Jones (ed.), Nebraska Symposium on Motivation , 3, 149–189. Lincoln, NE: University of Nebraska Press.

Petty, R.E. and Cacioppo, J.T. (1986) Communication and Persuasion: Central and Peripheral Routes to Attitude Change . New York: Springer Verlag. <http://dx.doi.org/10.1007/978-1-4612-4964-1>

Randall, D.M. and Wolff, J.A. (1994) The time interval in the intention-behaviour relationship: Meta-analysis . British Journal of Social Psychology , 33, 405–418. <http://dx.doi.org/10.1111/j.2044-8309.1994.tb01037.x>

Rimal, R.N. and Real, K. (2003) Understanding the influence of perceived norms on behaviors . *Communication Theory* , 13, 184–203.

Rivis, A. and Sheeran, P. (2003) Social influences and the theory of planned behaviour: Evidence for a direct relationship between prototypes and young people's exercise behaviour . *Psychology and Health* , 18, 567–583. <http://dx.doi.org/10.1080/0887044032000069883>

Rosenberg, M.J. (1956) Cognitive structure and attitudinal affect . *Journal of Abnormal and Social Psychology* , 53, 367–372. <http://dx.doi.org/10.1037/h0044579>

Rosenberg, M.J. (1965) Inconsistency arousal and reduction in attitude change . In I.D. Steiner, ed. and M. Fishbein (eds), *Current Studies in Social Psychology* , pp. 121–134. New York: Holt, Reinhart, and Winston.

Rosenberg, M.J. and Hovland, C.I. (1960) Cognitive, affective, and behavioral components of attitudes . In C.I. Hovland, ed. and M.J. Rosenberg (eds), *Attitude Organization and Change: An Analysis of Consistency Among Attitude Components* , pp. 1–14. New Haven, CT: Yale University Press.

Rosenstock, I.M., Strecher, V.J. and Becker, M.H. (1994) The health belief model and HIV risk behavior change . In R.J. DiClemente, ed. and J.L. Peterson (eds), *Preventing AIDS: Theories and Methods of Behavioral Interventions*. AIDS Prevention and Mental Health , pp. 5–24. New York: Plenum Press.

Schulze, R. and Wittmann, W.W. (2003) A meta-analysis of the theory of reasoned action and the theory of planned behavior: The principle of compatibility and multidimensionality of beliefs as moderators . In R. Schulze, ed. , H. Holling, ed. and D. Böhning (eds), *Meta-analysis: New Developments and Applications in Medical and Social Sciences* , pp. 219–250. Cambridge, MA: Hogrefe & Huber.

Sheeran, P. (2002) Intention-behavior relations: A conceptual and empirical review . In W. Stroebe, ed. and M. Hewstone (eds), *European Review of Social Psychology* , 12, 1–36. Chichester: Wiley. <http://dx.doi.org/10.1080/14792772143000003>

Sheeran, P. and Taylor, S. (1999) Predicting intentions to use condoms: A meta-analysis and comparison of the theories of reasoned action and planned behavior . Journal of Applied Social Psychology , 29, 1624–1675. <http://dx.doi.org/10.1111/j.1559-1816.1999.tb02045.x>

Sheppard, B.H., Hartwick, J. and Warshaw, P.R. (1988) The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research . Journal of Consumer Research , 15, 325–342. <http://dx.doi.org/10.1086/209170>

Srull, T.K. and Wyer, R.S. (1979) The role of category accessibility in the interpretation of information about persons: Some determinants and implications . Journal of Personality and Social Psychology , 37, 1660–1672. <http://dx.doi.org/10.1037/0022-3514.37.10.1660>

Triandis, H.C. (1972). The Analysis of Subjective Culture . New York: Wiley.

Triandis, H.C. (1977) Interpersonal Behavior . Monterey, CA: Brooks/Cole.

Tversky, A. and Kahneman, D. (1974) Judgment under uncertainty: Heuristics and biases . Science , 185, 1124–1131. <http://dx.doi.org/10.1126/science.185.4157.1124>

Uhlmann, E. and Swanson, J. (2004) Exposure to violent video games increases automatic aggressiveness . Journal of Adolescence , 27, 41–52. <http://dx.doi.org/10.1016/j.adolescence.2003.10.004>

Van Ryn, M. and Vinokur, A.D. (1992) How did it work? An examination of the mechanisms through which an intervention for the unemployed promoted job-search behavior . American Journal of Community Psychology , 20, 577–597.

Vroom, V.H. (1964) Work and Motivation . New York: Wiley.

Webb, T.L. and Sheeran, P. (2006) Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence . Psychological Bulletin , 132, 249–268. <http://dx.doi.org/10.1037/0033-2909.132.2.249>

- Wegner, D.M. (2002) *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
- Wegner, D.M. and Bargh, J.A. (1998) Control and automaticity in social life . In D.T. Gilbert, ed. , S.T. Fiske, ed. and L. Gardner (eds), *The Handbook of Social Psychology* Vol. 1 , 4th Edition, pp. 446–496. Boston, MA: McGraw-Hill.
- Wegner, D.M. and Wheatley, T. (1999) Apparent mental causation: Sources of the experience of will . *American Psychologist* , 54, 480–492. <http://dx.doi.org/10.1037/0003-066X.54.7.480>
- Weigel, R.H. and Newman, L.S. (1976) Increasing attitude-behavior correspondence by broadening the scope of the behavioral measure . *Journal of Personality and Social Psychology* , 33, 793–802. <http://dx.doi.org/10.1037/0022-3514.33.6.793>
- Weinberg, R.S., Gould, D., Yukelson, D. and Jackson, A. (1981) The effect of preexisting and manipulated self-efficacy on a competitive muscular endurance task . *Journal of Sport Psychology* , 3, 345–354.
- Wicker, A.W. (1969) Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects . *Journal of Social Issues* , 25, 41–78. <http://dx.doi.org/10.1111/j.1540-4560.1969.tb00619.x>
- Yang-Wallentin, F., Schmidt, P., Davidov, E. and Bamberg, S. (2004) Is there any interaction effect between intention and perceived behavioral control? *Methods of Psychological Research Online* , 8, 127–157.
<http://dx.doi.org/10.4135/9781446249215.n22>

Handbook of Theories of Social Psychology: Volume 1

Social Comparison Theory

Contributors: Jerry Suls & Ladd Wheeler

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Social Comparison Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n23>

Print pages: 460-483

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n23>

[p. 460 ↓]

Chapter 22: Social Comparison Theory

Jerry Suls, ed. and Ladd Wheeler, ed.

Abstract

This chapter reviews the history and development of the study of social comparison processes – the search for and utilization of information about other persons' standings and opinions for the purpose of accurate self-assessment or for self-esteem enhancement/protection. We describe the origins of Festinger's classic comparison theory in the group dynamics tradition, ambiguities in the classic theory, the later attributional reformulation of the theory, the research shift from self-evaluation to self-enhancement as a dominant motive for comparison (i.e., downward comparison theory [DCT]) to more recent social cognitive approaches which clarify what Festinger really meant (the proxy model) and emphasize the importance of knowledge accessibility (selective accessibility model) and social judgment (interpretation-comparison model) for self-evaluation. The essay concludes with some unresolved questions and illustrative applications of social comparison for education, health and subjective wellbeing.

Introduction

Although the statement, "Everything is relative," is an exaggeration, many things are relative, thereby complicating judgment and action. Apropos of a joke:

A snail was mugged by two turtles. When the police asked him what happened, he said: I don't know. It all happened so fast.

(Cathcart and Klein, 2007: 273)

Social philosophers and early social scientists recognized that people often rely on how they stand relative to other people to assess their opinions and potential. Leon Festinger (1954a) was the first to systematically consider this topic in "A theory of social comparison processes." "Social comparison" referred to the search for and utilization of information about other persons' standings and opinions for the purpose of self-assessment – judging the correctness of one's opinions, beliefs, and capabilities (Wood, 1996). Later researchers also placed emphasis on additional motives for comparison.

In this chapter, we survey the path of social comparison research from the 1950s to the current day. Although proceeding in chronological fashion, readers will see that the path was not straight, with some wrong [p. 461 ↓] turns, rest stops, and blind canyons. One of us (Ladd Wheeler) started on the trail in the early 1960s; the other (Jerry Suls) joined in the early 1970s. For two decades, we were "fellow travelers," but in the mid 90s we joined forces, a collaboration that has happily continued to the present. (Lefty and Dusty, the two cowboys, from *Prairie Home Companion* come to mind.) Because we are contributors to this research area, we may be biased in some conclusions so we will try to label our opinions as such. However, one lesson from social comparison theory is that distinguishing facts from opinions is no simple matter.

Group Dynamics

In the 1940s, Festinger was busy with research on informal communication in small groups at the Research Center for Group Dynamics at MIT and later at the University of Michigan. That work culminated in his theory of informal social communication which posited that people desire to attain uniformity of opinion either because group consensus provides confidence in one's opinion or because agreement was needed to achieve group goals (Festinger, 1950). The research with his students showed that patterns of communication and rejection of opinion deviates followed the propositions of the theory and provided a foundation for further study in conformity and group performance (e.g., Allen, 1965; Turner, 1991).

This group research prompted the Ford Foundation to award Festinger a grant to summarize and integrate empirical studies on social influence. In the 1950s, there was a keen interest to find ways to use social science research to inform policy and societal change. Festinger also was invited to give a talk at the second annual symposium on motivation held at the University of Nebraska. The talk was published as a chapter in an edited collection and “A theory of social comparison processes” was published in the journal, *Human Relations*, along with some related empirical papers, in the same year, 1954.

Both papers extended the arguments of informal communication theory, but the emphasis changed from the power of the group over the individual to how individuals use groups to evaluate the need to evaluate opinions and abilities. We now turn to details of the theory.

The Classic 1954 Formulation

Festinger began with the premise that people have the need to evaluate their abilities and opinions to be able to act in the world. He noted that they prefer to evaluate themselves relative to objective or physical standards, but these are not always available and in some cases never are. We can verify the Lone Ranger's (Clayton Moore) date of birth or whether the ground is muddy, but there are no objective standards for opinions about gun control or whether a cowboy will win a contest at the rodeo. But in many cases, we must make a judgment in the absence of objective information. In those instances, Festinger proposed that uncertainty induced a drive-like state that could be satisfied by social comparison (Recall in the 1950s drive theory was still popular in behavioral psychology).

A third key element in the theory was that comparison was important for the evaluation of opinions *and* abilities. The latter were not part of informal communication theory, but he realized, as noted above, that often there is no objective standard to assess our abilities, especially about future performance. Since abilities, unlike opinions, cannot be changed through communication, Festinger replaced “comparison” for “communication” – the prominent element in the earlier theory.

Festinger also recognized a distinctive feature of abilities. People want to be slightly better than everyone else because the desire to be better or to improve is emphasized in [p. 462 ↓] Western cultures. This tendency was referred to as the *unidirectional drive upward*.

These ideas were well received. The remainder of his *Human Relations* paper (1954a), however, focused on “the need for similarity” which might well have been the subtitle of the paper. We can evaluate our abilities and opinions *accurately* (1954a: 120; italics in original) only by comparison with other people of similar abilities and opinions. Consequently, we choose to compare to similar others, we try to change others to become more similar to ourselves, we change ourselves to be more similar to others, and failing all else, we simply cease comparing with dissimilar others. Why? Festinger claimed because we will not be able to make a “subjectively precise evaluation.” The startling aspect of the paper is that it never explains why similarity is necessary for accurate evaluation. However, this gap was not apparent for some time (see Deutsch and Krauss, 1965). Perhaps this is because the general idea that social comparison is a core element of social life seemed both evident and powerful.

Let's look at opinions and abilities separately to try to determine why similarity might be so important. In the case of opinions, Festinger wrote: “..., a person who believes that Negroes are the intellectual equals of whites does not evaluate his opinion by comparison with the opinion of a person who belongs to some very anti-Negro group” (1954a: 120–121). That comparison would lead to disagreement, and the equality believer would still not know if his opinion was correct. Instead we should compare with someone who is similar. If such a person agreed about the intellectual equality of Negroes and whites, the equality believer would be somewhat more confident in the correctness of the opinion. We are not told in which ways the comparison other should be similar beyond the fact that the person should not belong to a very anti-Negro group. Presumably, the person should be similar in values and world outlook, and we should have an expectation of agreement.

Turning to abilities, Festinger wrote: “..., a college student does not compare himself to inmates of an institution for the feeble minded to evaluate his own intelligence. Nor does a person who is just beginning to learn the game of chess compare himself to the recognized masters of the game” (1954a: 120). It is not clear how comparing to another

chess beginner would give one an accurate appraisal of one's chess ability. In fact, it would seem that comparing to a chess master would be more informative; one would at least have an idea of the level of ability that could be achieved. Bumbling around the board with another bumbler would say little about one's ability.

In short, Festinger's *Human Relations* article makes some degree of sense about the importance of similarity in feeling that our opinions are correct, but the argument regarding ability evaluation is completely unsatisfactory. There is no reason given about why we would be more accurate in ability appraisal as a result of comparing to someone of similar ability. Fortunately, the *Nebraska Symposium on Motivation* paper (1954b), written after the *Human Relations* article, is less formal, and offers a relatively clear explanation for the importance of similarity in ability evaluation. The explanation stresses action possibilities in the real world. We should be able to do the same things as others of similar ability can do. Below is a quote from the Nebraska paper illustrating this point.

Let us take as an illustration a person who tries to find out precisely how intelligent he is. Let us remember that his motivation is to know what his ability does and does not permit him to do in the real world in which he lives and acts. Simply knowing his score on some intelligence test does not tell him this. It does, however, allow him to compare himself with others. Suppose the others with whom he compares himself are all very divergent from himself in their score. He then knows that his possibilities for action in the world are very different from theirs. But this is negative knowledge and he still does not know precisely what he himself can do. Consider, however, the case where the others with whom he compares himself have scores very close [p. 463 ↓] to his own. He then knows that his own possibilities for action in the environment are identical or very similar to those for these other persons. This gives him the subjective feeling of knowing what he can or cannot do in the same way that having other agree with one's opinion gives the person a subjective feeling that his own opinion is correct.

1954b: 196–197

There are numerous other statements in the *Nebraska* chapter consistent with the one just given. For example:

[L]et us imagine a high school student who wants to know whether his intellectual ability is such as to enable him to go through college. He ... cannot adequately evaluate his ability in the real world. Clearly, going to college and seeing what happens would give him an evaluation of his ability for that purpose, but it is not possible to 'reality test' his ability before going to college. In the case of abilities it is particularly true that, even when a clear and unambiguous 'performance score' is available, it may not provide a satisfactory evaluation because of the large variety of situations for which the ability is relevant and the large number of purposes which make the ability important. Also important is the frequent desire to evaluate the ability *before* engaging in the action which would test it.

(1954b: 195; italics in original)

Thus, by emphasizing the importance of action possibilities in the *Nebraska* paper, Festinger sent a different message than in the *Human Relations* paper. The only part of the *Human Relations* paper that suggested that similarity is necessary to predicting what we can and cannot do is the Corollary to Derivation E that says:

An increase in the importance of an ability or an opinion, or an increase in its relevance to immediate behavior, will increase the drive toward reducing the discrepancies concerning that opinion or ability.

(1954a: 130¹)

We do know that almost everyone who subsequently worked in social comparison theory in the late 1950s to the mid 1970s used the *Human Relations* paper rather than the *Nebraska* chapter as their guide, and thus they simply accepted that similarity was important without knowing why it was important. The only paper that investigated the action possibility argument was by Steve Jones and Dennis Regan (1974) – a paper we will return to later.

Comparison Extended to Affiliation and Emotion

During the time Festinger was working on comparison theory, he already had formulated his theory of cognitive dissonance which was soon to become his dominant interest. We don't know whether Festinger recognized the ambiguity in comparison theory because he never returned to the topic.

Comparison theory might have languished, but for the use his ex-student Stanley Schachter made of it in his research on affiliation. In a series of clever experiments, Schachter (1959) showed that inducing fear in experimental subjects prompts them to want to wait with others, particular others who were also awaiting the same fearful stimulus (i.e., painful electric shocks) – inspiring the memorable, “Misery doesn't love just any company; misery loves miserable company.” Schachter thought that subjects preferred to affiliate to learn whether their emotional responses were appropriate for the circumstances. The fact that their preferred affiliates were people in the *same* circumstances (also waiting to receive the shocks) seemed to support Festinger's similarity hypothesis. (Subjects didn't want to wait with people who were there for some other purpose, such as waiting to see their advisors.)

The research had an enormous influence on the psychology of emotion, but gave little attention to the process of emotion comparison. As far as comparison studies were concerned, the biggest contribution was the recognition that comparisons apply to emotions as well as to abilities and opinions. This opened the door to apply social comparison to other attributes of the person in addition to abilities and opinions.

[p. 464 ↓]

Schachter's Indirect Effect and the Jesp Supplement

While Schachter pursued a different research direction, he taught a graduate course in group dynamics at the University of Minnesota that spread the gospel about informal communication and social comparison. Many of the graduate students, including Radloff, Hakmiller, Singer, Latané, Arrowood, and one of the present authors (Ladd Wheeler), conducted experiments on social comparison. In 1966, these were eventually published in a special supplement to the *Journal of Experimental Social Psychology*, Studies in Social Comparison, edited by Bibb Latané.

Space prevents us from a complete description of the supplement's contents, but the empirical findings showed support for some of the theory's propositions and extended its scope. People under stress don't just prefer to affiliate with others; they also grow to like others who are present, presumably because they satisfy the need for comparison (Latané et al., 1966). People prefer to affiliate with others who agree with them when uncertain about the correctness of their opinions (Gordon, 1966). Lacking comparison information, people are unstable and feel uncertain in their personal performance evaluations (Radloff, 1966). This study was what initially piqued JS's interest in comparison theory. Radloff extrapolated that exceptionally talented children might actually underperform academically because they lacked any similar peers with whom to compare. Similar others in these experiments were operationalized as "people in the same circumstances" or in agreement. There was little acknowledgment of the aforementioned ambiguity of the similarity concept.

This is most apparent, with hindsight, in the two most influential experiments in the supplement. Wheeler (1966) developed a procedure – subsequently known as the rank-order paradigm – to test the unidirectional drive upward idea which Festinger proposed would motivate the person to be slightly better than others (Only *slightly better*, according to Festinger, because the desire to be better was counteracted by the presumed desire to be similar). In Wheeler's experimental paradigm, subjects received a score and their rank on a bogus trait inventory. They were then offered the choice to see the actual score of someone else of a different rank to assess how they had done.

When subjects were motivated to think they were good, the likelihood of choosing to see the score of someone ranked above them (i.e., upward comparison) was greater. It is worth noting that those who chose to see the score of a higher ranked subject – in essence made an *upward comparison* – felt they were more similar to the person above them in the rank order than to the person below them. “The comparer is attempting to prove to himself that he is almost as good as the very good ones” (Wheeler, 1966: 30).

The second influential experiment was conducted by Hakmiller (1966), who was a graduate student with Wheeler and who adapted the rank order paradigm to introduce the idea of downward comparison. In the experiment, subjects were led to believe through a cover story that they received a high score on a supposed measure of “hostility toward one's parents.” In one condition, subjects were led to believe the high score represented a bad thing (high threat); in the other condition, they were led to believe it represented something relatively positive (low threat). Then all subjects were offered the opportunity to see someone else's score in the rank ordering. Threatened subjects were more interested in seeing the score of someone higher in hostility and therefore presumably worse off. This suggested to Hakmiller (1966) that threat to self-esteem increases the motivation to protect or enhance the self and thereby prompts downward comparison. The idea was a major source of inspiration for Wills' (1981) subsequent downward comparison theory that was influential a decade later.

[p. 465 ↓] The Wheeler (1966) and Hakmiller (1966) experiments had interesting implications – the former was suggestive of assimilative-type processes that later became central in social cognition research; the latter provided evidence for the motive to enhance or protect self-esteem. But, at the time, their most salient implication was that similarity should be conceived in terms of relative standing on the dimension under evaluation. In this way, similarity referred to people who shared the same opinions and performed in the same way. There is circularity in this, however, because in real life choosing to compare with a similar other presupposes that an implicit social comparison has already occurred (i.e., how does one know they are similar in the first place?). The rank-order paradigm solved this experimentally by providing rank information prior to the comparison choice. It took some time before the full extent of this problem was recognized.

Resting around the Campfire

Between the publication of the *JESP* supplement and the late 1970s, comparison was not a dominant interest of social psychology; the hot areas were cognitive dissonance and attribution theory. However, there were a few developments brewing. Morse and Gergen (1970) conducted an experiment to show that casual exposure to another person may be sufficient to produce an impact on a person's self-esteem. Job applicants encountered an accomplice whose personal appearance was either highly desirable (Mr. Clean) or highly undesirable (Mr. Dirty). Under the guise of collecting information about the applicants, subjects completed self-esteem scales both before and after exposure to Mr. Clean/Mr. Dirty. The results showed that exposure to Mr. Clean produced a decrease in self-esteem and exposure to Mr. Dirty produced an increase in self-esteem. "As a result of others' characteristics appearing more desirable or less desirable than his own, a person's generalized self-estimate is displaced downward or upward" (1970: 154). The Morse and Gergen experiment was important because it showed that social comparison was relevant to broader domains than specific abilities, opinions or emotions. Also, prior to this time, social comparison was considered to fall under the umbrella of "social influence," consistent with its origins in group dynamics. Morse and Gergen's research placed comparison under the umbrella of the psychology of the self.

Comparison's First Book

Comparison theory seemed to be at rest. There was research activity but studies were published piecemeal, there was no "big picture," and not even much of an invisible comparison college. (A candidate for the "big picture" was Tom Pettigrew's chapter, "Self-evaluation theories," in the 1967 *Nebraska Symposium on Motivation*, where he described the common themes and implications of social comparison, equity, and relative deprivation theories. But his chapter appeared in the same volume with Harold Kelley's "Attribution principles in social psychology," which clearly took most social psychologists' hearts and minds at that time.)

But primed by chutzpah, Radloff's experiments, Pettigrew's essay and the threat of unemployment more than any editing experience, JS and a colleague, Rick Miller, floated the idea of an edited collection of original essays on social comparison. To our surprise, we learned there were many social psychologists, some well known, who were conducting comparison research and who were eager to contribute. Persuading a publisher to take on this project was much harder than getting commitments from authors. At least one anonymous reviewer, described to us only as "an eminent social psychologist," told a prospective publisher that nothing new in social comparison had happened since 1959 [p. 466 ↓] so there was no need for a new book (JS thinks this was Schachter). But although publisher, Larry Erlbaum, did not offer us a contract, he persuaded another publisher (Hemisphere) to do so – a favor for which we remain grateful.

That volume, *Social Comparison Processes: Theoretical and Empirical Perspectives* (Suls and Miller, 1977), had 13 chapters on a broad range of comparison-related topics by both veteran and relatively junior social psychologists and a concluding commentary. Two chapters proved to be very influential. In "Pleasure and pain of social comparison," Brickman and Bulman proposed, counter to Festinger, that sometimes people want to vigorously avoid social comparison because of the threats it can pose. Their essay emphasized self-enhancement and self-protection rather self-evaluation. Brickman and Bulman's (1977) essay was a direct precursor of Thomas Ashby Wills' (1981) later downward comparison principles article. As a historical aside, Brickman was approached to write an update of his essay "Hedonic relativism and planning the good society," coauthored with Donald Campbell (1971). But in our conversations, Phil wanted us to reprint the original essay which he felt was little read (it is now considered a classic essay). JS recalls insisting that we could not reprint; we needed a new essay. Phil reluctantly agreed. Months later, he sent something entirely different that included six experiments specifically designed and conducted for the chapter.

The other chapter by Goethals and Darley (1977) received widespread acceptance when the volume was published. Ladd Wheeler, who coauthored a commentary with Miron Zuckerman for the book, thought this was the flagship chapter in the volume. (The commentary represented a change in plans. Originally, Jerry Suls and Miller were supposed to coauthor the commentary chapter, but Miller relocated to Germany and Jerry Suls had [realistic] doubts he could do it justice by himself. Ladd Wheeler

originally had been contracted to summarize the rank-order paradigm experiments, but he discovered it had too much overlap with another contributing author so he was invited to do the commentary. LW asked Miron Zuckerman to help him because LW had broken his writing arm sliding into second base during a softball game [not falling off a horse]).

Goethals and Darley proposed to resolve the ambiguity of the similarity hypothesis one had to consider the important role of related attributes. Related attributes are related to, and predictive of, an ability or an opinion. Festinger's Hypothesis VIII was: 'If persons who are very divergent from one's own opinion or ability are perceived as different from oneself on *attributes consistent with the divergence*, the tendency to narrow the range of comparability becomes stronger' (1954a: 133). Festinger noted that in level of aspiration studies, college students do not compete with high-scoring graduate students, because the college students see themselves as inferior on attributes consistent with the difference in scores. Goethals and Darley (1977) married Hypothesis VIII to attribution theory (Kelley, 1967, 1973) and argued that people will discount differences in performance if they are consistent with differences in related attributes. According to Kelley, the discounting principle is as follows: the role of a given cause in producing an effect is discounted if other plausible causes are also present.

If we want to evaluate an ability, we must observe a performance and make inferences about the underlying ability. There are many determinants of performance that are not indicative of the underlying ability – effort, luck, difficulty, age, practice, and so forth. Only if we compare our performance to that of someone similar on all these related attributes can we make a reasonable inference about our ability relative to the ability of the other person. This, according to Goethals and Darley, is what Festinger meant when he wrote that we must compare ourselves to similar others in order to make a subjectively precise self-evaluation.

[p. 467 ↓] The argument for the evaluation of opinions is somewhat different. Goethals and Darley distinguished between beliefs (a potentially verifiable assertion about the true nature of an entity) and values (liking or disliking of an entity). Only beliefs can be correct or incorrect, and they seem to correspond better than values to what Festinger meant by opinions. For that reason, we will discuss only beliefs. According to Kelley's attribution theory, a person needs to determine if his belief is entity-caused

(which means that he is correct) or person-caused (which means that his belief is a biased view based on his needs, values, wishes, etc). Goethals and Darley argued that comparison with others similar on related attributes would not be very useful unless they disagreed with the individual, in which case his belief in the correctness of his opinion would decrease. If similar others agreed, the individual would not know if their agreement was entity-caused or person-caused. If similar others disagreed, the individual would conclude that the disagreement was entity-caused. The same argument applies, *mutatis mutandis*, to comparison with others who are dissimilar on related attributes. Their disagreement would not be informative, but their agreement should increase one's confidence in the belief. Goethals and Darley concluded that they couldn't make a prediction about who would be chosen for comparison on the basis of related attributes.

We now view the related attributes hypothesis with less enthusiasm than when it appeared in 1977, but we still think it is a useful formulation for some purposes. If we compare our ability with someone who is perfectly similar to us on all related attributes, we can conclude that we are about as good as we ought to be, or better or worse than we ought to be, relative to that comparison person. That knowledge does not tell us, however, where we are on the ability scale, and it does not tell us what we can accomplish in the world. It is useful in helping us to evaluate a particular performance, which is often an important question. And if we have additional information about what the comparison person has accomplished in the world, we may be able to predict our own accomplishments. However, it was two decades before we fully appreciated these things and presented them in the "proxy model," to be described later (Suls et al., 2002).

Self-Enhancement and Downward Comparison

In the 1980s, Goethals and Darley were perceived to have resolved the ambiguity in Festinger's theory (Arrowood, 1986; Suls, 1986; Wood, 1989) and a series of experiments provided empirical support for the attributional reformulation (Gastorf and Suls, 1978; Suls et al., 1978; Wheeler and Koestner, 1984; Wheeler et al., 1982; Zanna et al., 1975). With self-evaluation apparently covered, the field moved on

to self-enhancement. Inspired in part by Hakmiller's experiment and Brickman and Bulmann (1977), Tom Wills (1981) presented an integrative review of prior research and original theory about how comparisons operate when the self is threatened. His paper, "Downward comparison principles in social psychology," covered a range of evidence from gossip and aggression to projection and self-esteem, arguing that downward comparison processes had a broad reach. For the most part, the studies Wills reviewed had been conducted to test other theories, but his analysis was persuasive. His review seemed to support two central ideas: threatened people (and those of low self-esteem) are more likely to compare with others who are worse off than themselves than do others who are better off (or of high self-esteem); and exposure to a less fortunate other (i.e., a downward target) boosts subjective wellbeing. DCT was broad, bold, and focused on "hot cognition" and it is not surprising that researchers found it appealing. For some time, a motive for self-enhancement, rather than self-evaluation, occupied the minds (and ambitions) of social psychological researchers.

[p. 468 ↓] DCT really took off due to a field study published by Joanne Wood, Shelley Taylor and Rosemary Lichtman (1985). They interviewed breast cancer survivors to learn how they adjusted to their condition. To the surprise of the researchers, a large majority of the women spontaneously reported coping better than other cancer patients (e.g., "I only had a lumpectomy, but those other women lost a breast.") In other words, the patients compared with others who were less fortunate. While Wills (1981) had relied on already-published studies, Wood et al. (1985) presented a new interesting result that reinforced his general thesis. Together, the papers showed that (downward) social comparison could serve as a means of coping, added to the novelty and applicability of DCT by making connections with the growing literature on stress and coping, health psychology, and the psychology of the self (Crocker and Major, 1989).

A stream of experiments followed showing how downward comparisons might aid the adjustment of medical patients and other stressed populations (see Gibbons and Gerrard, 1991; Tennen et al., 1991). Other researchers looked at psychological factors that moderated the motivation to compare downward and responses to such comparisons (Major et al., 1991; Tesser, 1998; Testa and Major, 1990). An entire edited volume by Bram Buunk and Rick Gibbons (1997) was devoted to coping and social comparison, mostly through downward comparison. DCT was a solid pillar of social psychology. Furthermore, by the 1990s a coherent narrative had emerged – people

seek out and are influenced by similar others on related attributes when they want to accurately assess their opinions and abilities; but they seek out downward comparisons when they want to reduce threat or enhance the self (Wood, 1989).²

Up and down

There were researchers, including ourselves, not as sanguine about the ubiquity or benefits of downward comparison. Hakmiller's (1966) threatened subjects did show more interest in seeing the score of someone higher in hostility, but most of the relevant evidence showed avoidance of more positive scorers (e.g., Smith and Insko, 1987). Reluctance to compare with the most fortunate is not the same as seeking someone worse-off (Wheeler, 2000; Wheeler and Miyake, 1992). Jerry Suls was not convinced because in several earlier experiments, he and Miller found that subjects who thought they had performed poorly still preferred to affiliate with people who had more ability (Suls and Miller, 1978). Of course, affiliation should not be equated with social comparison because wanting company can result from other motives. But Suls and Tesch (1978) found that even students who failed an exam preferred to know the high performance scores – scarcely evidence for downward comparison. However, we felt like party-poopers at the campfire – disconcerted about all of the fuss made about downward comparison. Truth be told, Dusty (Ladd Wheeler) was a confirmed nonbeliever, while Lefty (Jerry Suls) was more of an agnostic. We didn't doubt that downward comparison could make a person feel better, but would a threatened person be able to find and use downward targets?

The picture began to change as more evidence accumulated. First, there was the recognition that the prediction that downward comparison could protect threatened self-esteem was better supported by correlational survey and interview studies than experimental evidence. Also, contrary to downward comparison's prediction, people with high self-esteem seemed to benefit more from downward comparisons (Crocker et al., 1987). Another problem was that experiments that actually manipulated upward versus downward comparison often failed to include no-comparison control groups so it was unclear whether downward comparison increased well-being or upward comparison decreased it (Wheeler 2000). When LW measured spontaneous social comparisons in

[p. 469 ↓] daily life, he found people made downward comparison when they felt happy rather than unhappy – directly contrary to the DCT prediction (Wheeler and Miyake, 1992).

The solidity of downward comparison also was questioned by Taylor and Lobel (1989) who proposed that persons under threat might be served by both upward and downward comparisons. Perceiving oneself as better off than someone less fortunate can enhance wellbeing while having contact with someone more fortunate can provide inspiration and provide information about how to improve. The authors did not actually have any direct empirical evidence for these ideas, but as one of the authors, Shelley Taylor, was so strongly associated with the earlier breast cancer survivor downward comparison results and the paper was published in *Psychological Review*, the major theoretical journal in the field, the proposal received much notice.

The other article also included Taylor as a coauthor. Buunk et al. (1990) reported the results from two surveys that inquired about the degree to which cancer patients and college students experienced both positive and negative effects of making upward and downward comparisons. The results showed that direction of comparison was not intrinsically associated with positive or negative affect. In some cases, subjects reported positive reactions to upward comparison and negative reactions to downward comparisons. Affective responses seemed to depend less on the direction and more on the salient implication of the comparison – “Will I get better or will I get worse?” At the time, we had two opposing reactions to Buunk et al.’s article. On the positive side, we hoped the paper would give downward comparison researchers pause, but the argument rested on a correlational methodology and small effects.

Fortunately, experimental results eventually emerged showing that exposure to superstars produced positive effects on self-evaluations – a result directly in contradiction to downward comparison’s basic theses. Lockwood and Kunda (1997) had first- and fourth-year students read a newspaper article about an outstanding fourth-year student of matching major and gender. This student was outstanding in academics, student government, sports, and leadership. Subjects in the control condition were not exposed to a target. Afterward, subjects completed self-ratings. For fourth-year students, the superstar had no effect on self-ratings, but for the first-year students, exposure to the superstar was associated with higher self-ratings. The researchers

reasoned that the superstar could be inspiring to first-year students because such success was still attainable for them; this was, of course, not the case for fourth-year students. The idea that upward comparisons could be inspiring and could elicit positive affect led comparison researchers in a new direction. The analysis of the role of upward and downward comparison required a more nuanced treatment than offered heretofore.

The Return of Self-Evaluation: Proxy Rides into Town

During the heyday of downward comparison theory/research, one of the cowboys (Jerry Suls) still was obsessing about the attributional reformulation and whether it really resolved the original ambiguity in Festinger's theory. He had already begun to discuss this at some length with Renny Martin, a sharpshooter (the "Annie Oakley" of this story) at the University of Iowa. He also shared this worry with Ladd Wheeler at a Nags Head conference over a plate of barbecue and beans. Did comparison with someone similar in related attributes provide "an accurate evaluation" of one's ability, as Goethals and Darley (1977) claimed? If I outride someone who has the same experience on horseback, the same age, and so on (i.e., related attributes), what have I learned? It means I have more than met my potential. If the other person outrides me, then it means I have not met my potential. If both of us ride at the same level, then "I am as good as [p. 470 ↓] I ought to be." But this doesn't address whether I am a "good," "average," or "poor" rider. Goethals and Darley's analysis answered a perfectly fine question, but it wasn't the one they emphasized in their essay.

It was then that Lefty, Dusty and Annie reread Festinger's *Nebraska Symposium* chapter and realized that Festinger's question was "Can I do X?" and not "Do I perform X as well as I should?" We also looked at a paper published by Jones and Regan back in 1974, which was the only paper that investigated Festinger's action possibility argument. They showed, first, that people are most interested in comparative information about an ability when they anticipate making a decision about an action based on that ability. Second, they found that the preference for comparison with similar others is strongest when these similar others have experience using that ability in situations relevant to the participants' decisions.

Jones and Regan noted that

[w]hat appears to be a hidden assumption in the theory is that by comparison the individual learns not only where he stands relative to others but also something about what it means for his decisions, actions, and outcomes to stand at that level. In other words, the critical knowledge the individual seeks to gain concerns not so much the level of his ability as what he can and cannot accomplish with whatever level of ability he has.

(1974: 140–141)

We believe that Jones and Regan (1974) completely nailed social comparison theory. However, we don't recall reading the paper when it appeared in the *Journal of Experimental Social Psychology* (we would have read it), and the paper had minimal impact (four citations in the 70s and 80s). It is a nice illustration of scientists not being ready to hear something. When we finally came to accept the Jones and Regan (1974) argument, we extended it and developed what we called the “proxy model of social comparison” (Wheeler et al., 1997). Our basic assumption was that by comparing ourselves with a proxy who has attempted to perform a behavior X, we can determine our own likelihood of success at X. In order for this to be true, the proxy must be similar to us on the underlying ability. However, we can't directly observe ability, but only performance, and we need some way of knowing if the performance we observe is truly indicative of the proxy's ability.

Proxy extended the Jones and Regan (1974) argument by including the concept of “related attributes” from Goethals and Darley (1977) and recognizing the importance of maximal effort. A proxy's prior success on a novel task ('X') should be a good index of one's likely future performance on “X,” if both self and proxy performed similarly on a prior related task and the proxy is known to have exerted maximal effort on that occasion. If it is unclear whether proxy made a maximum effort then proxy may be an inappropriate comparison (e.g., if proxy was fatigued on the first performance, it may be an underestimate of what they can do and a poor prognosticator for one's own future success).

Another element of proxy theory is that if information about proxy's maximum effort is unavailable, then proxy's success/failure can still be informative if the individual and proxy share related attribute standing. Finally, knowing proxy made a maximum effort and performed the same as self on a prior (related) task means similarity on related attributes is not critical. This means that related attribute similarity is not always required for a proxy to be an appropriate and useful comparison other.

The “proxy” research was an intellectual satisfying collaborative effort for the three of us, resulting in a theoretical article (Wheeler et al., 1997) in the first issue of *Personality and Social Psychology Review*. The theory article was followed by a series of four experiments supporting the predictions described above (Martin et al., 2002) and publication of two chapters in edited collections (Martin, 2000; Suls et al., 2000). (We should note around the same time William Smith of Vanderbilt also recognized [p. 471 ↓] the limitations of the Goethals and Darley formulation and, in fact, was the first to use the term “proxy”; see Smith and Sachs, 1997).

The work on proxy was satisfying to us, but like Jones and Regan (1974), nobody cites it. Most writers continue to refer to Goethals and Darley (1977) without recognition that, like Festinger (1954a), their approach was ambiguous about what self-evaluative question it answers. A cynical cowboy might conclude that perhaps the unexamined life is worth living (or at least being cited for). More likely, the lack of interest is a result of current social psychological excitement about automatic processes, heuristics, and hot cognitions. Proxy is about deliberative processing and will probably remain uncited until the social psychology pendulum swings back, as we trust and hope it will (prior to our move to the Cowboy Retirement Home).

In the meanwhile, we continue to maintain that the questions subsumed under the “self-evaluation” umbrella need to carefully distinguished. On some occasions, people want to know, “Do I play golf as well as I ought to?” On other occasions, especially where there are significant costs associated with failing, they want to know, “Can I win this match?” There are probably also other kinds of self-evaluative questions waiting for consideration (for our “spin” on opinion comparison, see Suls et al., 2000).

Social Cognition's Turn

Although social comparison researchers had ingested attribution theory and the psychology of self in the 1970s to 1990s (Suls & Wheeler, 2000), the area showed little recognition or application of basic cognitive process of judgment and perception as Kruglanski and Mayseless (1990) and Wills and Suls (1991) noted. Social comparison should be relevant to and informed by what psychologists have learned about psychophysical processes (Herr et al., 1983; Parducci, 1974), judgments of similarity and dissimilarity (Tversky, 1977), automatic processing, and social cognition. This began to be recognized in the mid 1990s and happily continues to the present.

In a *Psychological Bulletin* review, Collins (1996) resurrected the Wheeler (1966) results for upward comparison. Her basic idea was that upward comparison could lead either to contrast or to assimilation, depending on whether the comparison is construed as indicating similarity to, or difference from, the comparison person. A major factor in this determination is the comparer's expectation, which acts as a top-down cognitive influence on perception and judgment. People are more likely to perceive similarity when they expect to find it (Manis et al., 1991). Since people want to believe they have positive characteristics, they perceive similarity with upward targets and conclude, "[T]hey are among the better ones" (Collins, 2000: 170). By this time, sufficient evidence was available demonstrating that social comparison could lead to contrast or assimilation. The other major lesson was that comparison outcomes are not intrinsically linked to the direction of the comparison (Buunk et al., 1990). It remained to be determined what were the critical factors leading an assimilative or contrastive outcome.

Two Contenders at the Comparison Corral

Once assimilation was advanced as a potential outcome of social comparison, connections to the priming literature in social cognition became apparent. The basic idea of priming is that increasing a category's cognitive accessibility, by prior use or subliminal presentation, increases the probability for a new ambiguous stimulus to

be interpreted consistently with the accessible category (Bruner, 1957; Neely, 1976, 1991). Priming a concept, for example, by having subjects unscramble four-word sequences to create three-word [p. 472 ↓] sentences that described behaviors related to a trait construct such as hostility, would cognitively activate the concept of hostility. Subjects who subsequently read a passage describing a target's activities ambiguous with respect to hostility (e.g., complained to a store clerk) would then rate the target's behaviors as more hostile (Srull and Wyer, 1979; also see Higgins, 1996; for a meta-analysis see DeCoster and Claypool, 2004). In essence, priming moves the evaluation of the target stimulus closer to the primed standard – assimilation. Subsequent research (Lombardi et al., 1987) showed that under certain conditions, priming also could displace the judgment of the stimulus away from the primed standard – contrast. Extension from other-evaluation to social comparison and self-evaluation was a logical development. Two major theories have emerged.

Selective Accessibility Model (SAM)

In Mussweiler's (2003; see also Mussweiler and Strack, 2000) approach, whether the response to social comparison elicits self-evaluations toward or away from the comparison target depends on what information is cognitively accessible. Almost at the moment of exposure, an initial holistic assessment of the similarity between the target and the self is made. This is consistent with cognitive psychology, which finds that people rapidly consider a small number of salient features to determine whether an object and a target are generally similar or dissimilar. In the case of social comparison, this might be salient characteristics, such as gender or age. Then, the holistic impression prompts subsequent information retrieval that focuses on hypothesis-consistent evidence (Klayman and Ha, 1987). A general impression of similarity between the comparer and the target sets in motion a process of "similarity testing." Because there are many facets of the self, people can construe self-knowledge in such a way that accessible knowledge is consistent with the initial holistic impression. The consequence is that the self-evaluations should draw closer to the target after selective search for similarity, leading to assimilation.

There are cases where the initial holistic impression is one of dissimilarity, which should prompt selective retrieval of target-inconsistent knowledge about the self, or construal

consistent with the initial impression of difference. After dissimilarity search, self-evaluations should be displaced from the comparison target, leading to contrast.

Empirical support for SAM is good (e.g., Mussweiler et al., 2005). For example, priming social standards makes certain kinds of standard-consistent or standard-inconsistent information more cognitively accessible and hence faster to access in lexical priming tasks (Mussweiler and Strack, 2000). In addition, if subjects are initially led to focus on similarities versus differences; for example, between two pictorial scenes, this primes subsequent similarity versus dissimilarity testing with respect to a comparison target, such as another college student described as adjusting well or poorly. When the subjects subsequently evaluated their own college adjustment, those who had been primed to focus on similarities listed more social activities and friends after comparison with a very adjusted target than after a poorly adjusted target. Those subjects primed to focus on differences reported their adjustment to college was worse after comparison with a well-adjusted target than a poorly adjusted target. Assimilation (versus contrast) with a comparison target occurred depending on whether subjects had been primed to focus on similarities or differences.

SAM has the dual virtues of being parsimonious and explaining why several factors are moderators of assimilation/contrast. These factors include attainability (Lockwood and Kunda, 1997), perceived control (Major et al., 2001), and psychological closeness to the comparison target (Brown et al., 1992; Mussweiler and Bodenhausen, 2002). They are important because each should lead to a holistic impression of similarity, promoting further testing for similarity. If information [p. 473 ↓] about the self is retrieved that is consistent with the target then assimilation should result. Conversely, unattainability, low control, the absence of psychological closeness, or extreme comparison targets (Stapel and Koomen, 2000) should produce an initial impression of dissimilarity, thereby setting in motion a search for differences and eventuating in contrastive self-evaluations.

A distinctive feature of SAM is its contention that selective accessibility operates directly on the mental representations of self-evaluation. This contrasts with psycho-physicists (Biernat et al., 1991) who claim that a comparison target may serve as an anchor that changes the meaning of, and interpretation of, the judgmental scale. In the company of a professional basketball player, two cowboys might rate themselves as “very short,”

though objectively, the cowboys might be of average height. The “pro” might serve as a reference point that influences the interpretation of the points on the rating scale so that “average” and “short” take on a different meaning.

To avoid the possibility that it is merely the language or ratings people use to make comparative judgments, Mussweiler (2003) prefers to have subjects make absolute judgments rather than subjective ratings. Thus, instead of asking subjects how tall they are on a seven-point scale with (“very tall” to “very short”), he would ask how tall they are in feet and inches. The latter should represent the comparer’s underlying mental representation.

A final feature of SAM is that similarity testing is assumed to operate as a default. This is consistent with the literature on judgment and decision-making where people initially tend to focus on similarities rather than differences in comparisons (Chapman and Johnson, 1999). Of course, in SAM, the search for similarity is short-circuited when a target is initially perceived as extreme or distinct, does not share related attributes, or has some salient attribute suggesting the self and target are probably quite different. The holistic impression of dissimilarity should set in motion a search for differences that leads to contrast.

Interpretation-Comparison Model (ICOM)

The second approach (Stapel, 2007; Stapel and Koomen, 2000) also depends on knowledge accessibility but emphasizes the way social comparison information is *used*. ICOM posits that social comparison can instigate two different processes with opposing effects. The standing of a comparison target may provide an interpretative framework to define the self and be incorporated in the self-definition with assimilation (as priming hostility prompts the interpretation of a person’s ambiguous behaviors as aggressive). Alternatively, the same comparison may serve as an extreme standard against which the self is evaluated such that the information is excluded from the self-concept – producing contrast. Stapel (2007) thinks both assimilative and contrastive processes can occur simultaneously, but when the interpretative pull is stronger, assimilation is the consequence. When the comparative push is stronger, contrast results.

The degree of push versus pull depends on several factors. One example is the *distinctness* of the comparison. When the behavior of a comparison target activates distinct actor-trait links (“Gene Autry is a skilled horseman”), self-evaluation is likely to be contrastive (“I am a poor rider”); if the target activates indistinct trait information (“skilled horseman”), the outcome is assimilative (“I am a skilled horseman”). The latter outcome is especially likely when the comparer’s self-concept is mutable or unclear, such that there is room for inclusion and/or a need for filling in the gaps in the self-concept (Stapel and Koomen, 2000).

Another consideration is whether people are *explicitly* asked to compare themselves with a comparison standard (“Do you sing as well as Hank Williams?”) or whether the comparison is *implicit* (“Think about Hank Williams”). Stapel and Suls (2004) reasoned that intentionally comparing will make self-relevant information accessible that is [p. 474 ↓] consistent with the comparison target because a common ground is needed on which the target and comparer can be compared (Gentner and Markman, 1994). This similarity focus should be less likely to be a feature of spontaneous, effortless reactions to implicit social comparisons. Results showed that explicit social comparisons more often lead to assimilation with a superior or inferior comparison target; implicit comparisons were associated with contrastive self-evaluations. Explicit comparisons are also subject to more boundary conditions. For example, if the self is not perceived a mutable or there is too much delay between comparison and self-evaluation then contrast “wins” over assimilation. Contrast appears to be a more typical outcome. With hindsight, this might account for why the social comparison field took longer than expected to appreciate the possibility of assimilative outcomes.

It's a Draw for Now

Currently, Mussweiler’s and Stapel’s theories are the most popular and well-supported approaches for understanding responses to upward and comparison. While they share an underlying premise about comparison-induced knowledge accessibility, they differ in important aspects. For SAM, everything depends on whether similarity testing versus difference testing is initiated – an “either/or” process. ICOM assumes outcomes are driven by interpretation or comparison with a standard, but both may be simultaneously operating. The evaluative outcome depends on whether interpretation or comparison

“wins out.” SAM posits that similarity testing tends to be a cognitive default, while ICOM does not. SAM posits that absolute judgments and behaviors are the best proof of comparison-induced changes in self-evaluations. ICOM finds consistent evidence for subjective rating, absolute judgments, and behavioral outcomes and assumes the same underlying processes are responsible. SAM assumes the same processes and outcomes result whether comparisons are implicit or explicit; ICOM argues they have different effects. SAM seems more parsimonious to us, but only strong inference testing between the two theories will decide.

Definitions and Control Groups

Comparison research inspired by social cognition has focused on forced comparisons subject to, for the most part, automatic processes. We perceive progress although there are disparate results that could lead to the impression that there are no strong conclusions to be drawn. We don't share that impression, but we are keenly aware of sources of confusion and the need for better research.

One source of confusion is definitional. In a recent survey of the literature on assimilation in social comparison (Wheeler and Suls, 2007), we identified several kinds of outcomes measured in recent studies of assimilation-contrast: (single) attribute evaluations, global evaluation, mood, and behavior. Throughout this chapter, we have used the terms “contrast” and “assimilation” carefully to refer to self-evaluations displaced toward or away from the comparison target. The literature is not as consistent; there is a tendency for researchers to lump together as evidence different kinds of outcomes and not fully acknowledge nonparallel findings. For example, Buunk and Ybema (2003) exposed Dutch married women to a description by another married woman of her either happy or unhappy marriage. Affect was more positive following an upward comparison, but subjective evaluation of the participant's relationship was lower. The authors concluded that this demonstrated assimilation on affect and contrast on self-evaluation. According to our definition, assimilation can occur only on self-evaluation; it cannot occur on affect. There are various reasons why affect might increase following exposure to a successful other, but assimilation is not one of them.

In addition, some writers equate assimilation with Tesser's (1988) reflection [p. 475 ↓] or basking-in-reflected-glory. As we observed elsewhere, "feeling good because one knows Mick Jagger's friend is not the same as believing one can perform before crowds in Madison Square Garden," (Suls and Wheeler, 2007: 35). Assimilation simply does not exist in Tesser's model; all comparison results in contrast. The reflection process leads to positive affect but not to increased self-evaluation on the comparison dimension.

A second problem is the missing control group problem (Wheeler, 2000; Wheeler and Suls, 2007). Researchers rarely include no-comparison control groups in their designs so it is not possible to determine whether upward target or downward target is attracting or repelling or both are acting. This is a critical issue because inclusion of a control group is the only way to know whether the cognitive processes posited by Mussweiler and Stapel are actually operating as they claim. Also, both cognitive models imply that upward assimilation and downward assimilation are just as likely given the right circumstances. In our reviews of the literature, however, while we find that upward assimilation appears to be robust, downward assimilation is rare: people do not assimilate to those who are worse off. This is consistent with Collins' (1996) assertion that people generally think they are good and assume similarity upward, not downward.

Self-Enhancement

The latter observation provides a transition to the third issue – the role of the self-enhancement motive. It is a bit ironic that responses to upward and downward comparison were originally connected to self-enhancement, but currently popular theories, such as Stapel's and Mussweiler's, emphasize cognitive processes and cognitive moderators. To address this gap, our recent work has involved the development of an approach/avoidance model of upward-downward comparison effects (Suls and Wheeler, 2008). Our framework borrows heavily from SAM but recognizes that self-enhancement lumps together two very different motives – the need to excel and the need to avoid failure. Whether our approach-avoidance model properly integrates cognitive and affective processes only more time on the research trail will tell.

Applicability to Social Issues

This chapter has emphasized the progression of theory and evidence, but social comparison also has offered insights into a wide variety of social problems and applications. In the interests of space, our examples will be selective.

Education

Radloff (1966) found that extreme scorers, who, by definition, would have difficulty finding similar comparison others, are very uncertain about their abilities and fluctuate in their self-evaluations. He suggested this is a good reason to institute ability tracking in the schools – to provide both gifted and untalented classmates with appropriate peers with whom to compare. It has been documented, however, that tracking is associated with negative academic self-concept, especially in higher-ability classrooms. Students with the same ability (as measured by standardized tests) typically have *lower* academic self-concept when they attend *higher-ability* schools than when they attend lower-ability schools, a finding known as the *big-fish-little-pond effect* ("BFLPE"; Marsh and Hau, 2003). The BFLPE is explained as a negative contrast effect created by invidious classroom comparisons, but heretofore little direct evidence for comparisons mediating the negative contrast effect. The BFLPE also seems to be contradicted by consistent findings that students' performance improves if they report comparing their exam grades with a classmate who performed (slightly) *better* than [p. 476 ↓] themselves (Blanton et al., 1999; Huguet et al., 2001).

We (Wheeler and Suls, 2005) proposed that forced upward social comparisons, characteristic of BFLPE, and strategic deliberate upward comparison choices can coexist. Students in low- and high-ability schools may deliberately select classmates with *slightly better* grades (and therefore attainable accomplishments) as comparison targets, but students in high-ability schools are also involuntarily exposed to "superstars" (whose accomplishments might be seen as unattainable), and thus suffer a decline in self-concept. The net result would be a lower academic self-concept in the high-ability schools or classes.

With an international team of collaborators including Lefty and Dusty, Pascal Huguet and Florence Dumas tested the role of deliberate and forced comparisons and their effect on academic self-concept in a sample of more than 2,000 French secondary school students. Measures of academic self-concept and perceived classroom standing and nominations of the classmate with whom they preferred to compare their grades were collected. Students' aptitude test scores were also available. The negative contrast effect was replicated in this sample, but it disappeared after statistically controlling for students' invidious comparisons in the classroom, providing the first direct evidence that social comparison mediates the BFLPE. But the BFLPE coexisted with the tendency for students to nominate a better-performing classmate to compare their grades. Moreover, this tendency was positively related to academic self-concept, suggestive of an assimilation effect. After statistically controlling for the comparison choice effect, the magnitude of BFLPE was reduced, but not eliminated. Huguet et al. (2009) concluded that BFLPE is the net effect of counterbalancing influences: stronger negative contrast effects associated with forced exposure to invidious comparisons at the classroom level and weaker assimilation effects associated with upward social comparison choices. For educational policy makers, this work suggests that academic selective schools do not automatically benefit the students who attend them, contrary to a largely uncritical belief.

Health

The original application of downward comparison to the adaptation of the chronically ill medical patients was already described. This was mainly correlational-descriptive research (Buunk and Gibbons, 1997; Wood et al., 1985) and has not caught up with recent lab findings on forced automatic processes. An exception is Stanton et al. (1999) who assigned breast cancer survivors to listen to an audiotape interview of a (supposed) patient whose comments reflected good, poor, or unspecified psychological and physical health status. While patients who had listened to the poorly adjusted patient rated their own adjustment as better than those exposed to the well-adjusted patient, even those who heard a poor-functioning survivor reported feeling better than the patient in the interview. In short, and reminiscent of Buunk et al. (1990), patients were able to find positive meaning in either direction. Their findings also reinforce our

belief that downward assimilation rarely occurs, even when a downward comparison is forced. In any case, the kind of research conducted by Stanton et al. has the potential to inform medical professionals about what kinds of patient models and instructions are appropriate for psycho-educational materials, such as videos used in medical rehabilitation.

The relevance of proxy and upward comparison can be seen in Kulik and Mahler's (1987) demonstration that coronary-bypass surgery patients experience better adjustment post-bypass if prior to the procedure they are assigned to share a room with a patient who already undergone bypass (versus a patient who also is waiting). Assigning patients to different roommates is feasible and has implications for cardiology practice.

[p. 477 ↓] Support groups allowing for open disclosure and discussion are a standard psychosocial referral for cancer patients. These groups tend to include all patients with a particular cancer, regardless of the need for psychosocial treatment, which means the groups are heterogeneous in terms of distress level. One concern is that nondistressed members may actually suffer deficits in such groups, leading to suggestions that groups should be homogenous in composition. This idea has been resisted because nondistressed patients are thought to serve as important role models. Several reviews find that there are overall improvements in depression, anxiety, and self-efficacy as a function of group interventions (e.g. Spiegel et al., 1981), but the biggest gains are typically seen in those patients with the fewest psychosocial resources.

Support groups fulfill needs for affiliation, receipt of information, emotional ventilation, and social comparison. A recent social comparison analysis of the social support group process (Carmack Taylor et al., 2007) drew on findings from the upward and downward comparison literature. On this basis, they concluded that distressed patients need to be in heterogeneous-composition groups so they are exposed to better functioning patients from whom they can be inspired, learn, and assimilate. There is insufficient evidence to determine whether nondistressed patients may receive no benefits or even experience negative effects from associating with low-functioning patients. Thus far, only a single study has shown that patients with ample psychosocial resources may be harmed by the group experience (Helgeson et al., 2000). Based on our previous comments, we are doubtful that nondistressed patients experience downward assimilation, but they may

find the support experience irrelevant for them. Carmack Taylor et al. correctly call for more research, but propose that well-functioning patients might be recruited to serve as “group volunteers” and trained to meet the needs of distressed patients for affiliation and information. Framed in this way, nondistressed patients may desire participation and volunteer their time in the spirit of altruism.

Economics, Subjective Wellbeing, and Greed

We are writing this chapter during serious economic times. The stock market lost millions of dollars, housing mortgages are in default, and the federal government has given record amounts of money to banks to keep them afloat. Social comparisons are prevalent and irritating. One large company, which has already received billions of dollars in bailout money from the federal government and has asked for more, gave top company administrators million dollar bonuses. Letters to the editor from “average wage-earners” complained about the unfairness of the huge bonuses while more people are losing their jobs. It is obvious these highly paid CEO's are creating invidious comparisons that effect short-term changes in subjective wellbeing among the general public.

The current economic recession/depression has been explained by faulty regulation in the last two presidencies and increasing levels of greed. The reason for a lack of regulatory oversight is clear, but why should greed have increased? It has many sources, social comparison being one. This is very nicely demonstrated by an archival analysis tracking the causal impact of the introduction of television from 1951 to 1955 on FBI indicators of crime, burglary, auto theft, and larceny in the US (Hennigan et al., 1982). The researchers hypothesized that as the possessions and lifestyles of the advantaged become increasingly visible to the less advantaged, opportunities for comparison and frustration increase. Television allows everyone to learn about both the availability of goods and the conspicuous consumption of the “rich and famous.” In a time-series design, the introduction of television in different geographical regions had no effect on violent crimes, burglary, or auto theft. [p. 478 ↓] However, larceny increased as television was introduced. If one feels one deserves desirable goods, but

lack the legal means, then theft becomes an option. In recent years, exposure to luxury goods, vacations, and wealth has become even more prevalent – leading to invidious social comparisons and perhaps increasing greed. The current economic predicament provides a veritable comparison laboratory.

Is This the End of the Trail?

This chapter has covered more than 50 years of theorizing and research to tell the story of social comparison. In 40 pages, this could not be the complete story. There are more yarns and other cowboys and cowgirls to tell them. One thing we are sure of is that comparison, after a slow start, has shown sustained growth – we think because it is a core element of social life. The lesson may be to never kick what you think is a dead horse; it may only be getting a few winks of sleep while waiting for the sun to come-up and the cattle drive to continue.

Notes

1 The prose in this and other sections of the *Human Relations* article is tedious. Anyone who is familiar with Festinger's other publications knows he was capable of very strong writing. Someone once speculated that Festinger wrote the *Human Relations* paper in a style intended as a sly parody of Clark Hull and Kenneth Spence's theoretical papers which were filled with dense hypotheses and corollaries. (Festinger went to graduate school at the University of Iowa and for a time was Spence's research assistant.) This is such a good story and so consistent with Festinger's sense of fun that we hope it is true.

2 In this chapter, we focus on self-evaluation and self-enhancement as comparison motives. Wood (1989) suggested a third motive – self-improvement. This refers to studying the way experts do things (e.g., swing a golf club, play a piano) in order to imitate them or by an upward comparison. We are reluctant to accept a need for a third motive, however. First, how can we distinguish inspiration provided by a role model and self-evaluation (with unidirectional drive)? Second, self-improvement simply may be a form of self-enhancement because by improving, self-concept should be enhanced.

References

- Allen, V. (1965) Situational factors in conformity . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 2, 133–175. New York: Academic Press.
- Arrowood, A.J. (1986) Comments on 'Social comparison theory: Psychology from the lost and found' . *Personality and Social Psychology Bulletin* , 12, 279–281. <http://dx.doi.org/10.1177/0146167286123002>
- Blanton, H., Buunk, B.P., Gibbons, F.X. and Kuyper, H. (1999) When better-than-others compare upward: Choice of comparison and comparative evaluation as independent predictors of academic performance . *Journal of Personality and Social Psychology* , 76, 420–430. <http://dx.doi.org/10.1037/0022-3514.76.3.420>
- Brickman, P. and Bulman, R.J. (1977) Pleasure and pain in social comparison . In J. Suls, ed. and R.L. Miller (eds), *Social Comparison Processes: Theoretical and Empirical Perspectives* , pp. 149–186. Washington, DC: Hemisphere.
- Brickman, P. and Campbell, D.T. (1971) Hedonic relativism and planning the good society . In M.H. Appley (ed.). *Adaptation-level Theory: A Symposium* , pp. 287–302. New York: Academic Press.
- Brown, J.D., Novick, N.J., Lord, K.A. and Richards, J.M. (1992) When Gulliver travels: Social context, psychological closeness, and self-appraisals . *Journal of Personality and Social Psychology* , 62, 717–727. <http://dx.doi.org/10.1037/0022-3514.62.5.717>
- Bruner, J.S. (1957) On perceptual readiness . *Psychological Review* , 64, 123–152. <http://dx.doi.org/10.1037/h0043805>
- Buunk, B.P., Collins, R.L., Taylor, S.E., VanYoeren, N.W. and Dakof, G.A. (1990) The affective consequences of social comparison: Either direction has its ups and downs . *Journal of Personality and Social Psychology* , 59, 1238–1249. <http://dx.doi.org/10.1037/0022-3514.59.6.1238>

Buunk, B., ed. and Gibbons, F.X. (eds) (1997) *Health, Coping and Well-being: Perspectives from Social Comparison Theory*. Mahwah, NJ: Lawrence Erlbaum Associates.

Buunk, B.P. and Ybema, J.F. (2003) Feeling bad, but satisfied: The effect of upward and downward comparison upon mood and marital satisfaction . *British Journal of Social Psychology* , 42, 613–628. <http://dx.doi.org/10.1348/014466603322595301>

Carmack Taylor, C.L., Kulik, J., Badr, H., Smith, M., Basen-Engquist, K., Penedo, F. and Gritz, E.R. (2007) A social comparison theory analysis of group composition and efficacy of cancer support group programs . *Social Science and Medicine* , 65, 262–273. <http://dx.doi.org/10.1016/j.socscimed.2007.03.024>

Cathcart, T. and Klein, D. (2007) *Plato and a Platypus Walk into a Bar* . New York: Penguin Books.

Collins, R.L. (1996) For better or for worse: The impact of upward social comparisons on self-evaluations . *Psychological Bulletin* , 119, 51–69. <http://dx.doi.org/10.1037/0033-2909.119.1.51>

Collins, R.L. (2000) Among the better ones: Upward assimilation in social comparison . In J. Suls, ed. and L. Wheeler (eds), *Handbook of Social Comparison: Theory and Research* , pp. 159–171. New York: Kluwer Academic/Plenum. http://dx.doi.org/10.1007/978-1-4615-4237-7_9

Crocker, J. and Major, B. (1989) Social stigma and self-esteem: The self-protective properties of stigma . *Psychological Review* , 96, 608–630. <http://dx.doi.org/10.1037/0033-295X.96.4.608>

Crocker, J., Thompson, L., McGraw, K. and Ingberman, C. (1987) Downward comparison prejudice and evaluations of others: Effects of self-esteem and threat . *Journal of Personality and Social Psychology* , 52, 907–916. <http://dx.doi.org/10.1037/0022-3514.52.5.907>

DeCoster, J. and Claypool, H.M. (2004) A meta-analysis of priming effects on impression formation supporting a general model of information biases . *Personality and Social Psychology Review* , 8, 2–27. http://dx.doi.org/10.1207/S15327957PSPR0801_1

Deutsch, M. and Krauss, R.M. (1965) Theories in Social Psychology New York: Basic Books.

Festinger, L. (1950) Informal social communication theory . *Psychological Review* , 57, 271–282. <http://dx.doi.org/10.1037/h0056932>

Festinger, L. (1954a) A theory of social comparison processes . *Human Relations* , 7, 117–140. <http://dx.doi.org/10.1177/001872675400700202>

Festinger, L. (1954b) Motivation leading to social behavior . In M.R. Jones (ed.), *Nebraska Symposium on Motivation* , pp. 191–218. Lincoln, NE: University of Nebraska Press.

Gastorf, J.W. and Suls, J. (1978) Performance evaluation via social comparison: Related attribute similarity vs. performance similarity . *Social Psychology Quarterly* , 41, 297–305. <http://dx.doi.org/10.2307/3033582>

Gibbons, F.X. and Gerrard, M. (1991). Downward comparison and coping with threat. In J. Suls, ed. and T.A. Wills (eds), *Social Comparison: Contemporary Theory and Research* , pp. 317–346. Hillsdale, NJ: Lawrence Erlbaum.

Goethals, G. and Darley, J. (1977) Social comparison theory: An attributional approach . In J. Suls, ed. , and R. Miller (eds), *Social Comparison Processes: Theoretical and Empirical Perspectives* , pp. 259–278. Washington, DC: Hemisphere.

Gordon, B.F. (1966) Influence and social comparison as motives for affiliation . *Journal of Experimental Social Psychology*, Suppl. 1 , 2, 55–65. <http://dx.doi.org/10.1016/0022-1031%2866%2990066-7>

Helgeson, V.S., Cohen, S., Schulz, R. and Yasko, J. (2000) Group support interventions for women with breast cancer: Who benefits from what? *Health Psychology* , 19, 107–114. <http://dx.doi.org/10.1037/0278-6133.19.2.107>

Hennigan, K., Rosario, M., Heath, L., Cook, T.D., Wharton, J. and Calder, B. (1982) Impact of the introduction of television on crime in the United States: Empirical findings and theoretical implications . *Journal of Personality and Social Psychology* , 42, 461–477. <http://dx.doi.org/10.1037/0022-3514.42.3.461>

Herr, P.M., Sherman, S.J. and Fazio, R.H. (1983) On the consequences of priming: Assimilation and contrast effects . *Journal of Experimental Social Psychology* , 19, 323–340. <http://dx.doi.org/10.1016/0022-1031%2883%2990026-4>

Higgins, E.T. (1996) Knowledge activation: Accessibility, applicability and salience . In E.T. Higgins, ed. and A. Kruglanski (eds), *Social Psychology: Handbook of Basic Principles* , pp. 133–168. New York: Guilford Press.

Huguet, P., Dumas, F., Marsh. H., Seaton, M., Wheeler. L., Suls, J., Regner, I. and Nezlek, J. (2009) Clarifying the role of social comparison in the big-fish-little-pond effect (BFLPE): An integrative study . *Journal of Personality and Social Psychology: Interpersonal Relationships and Group Processes* . 97, 156–170. <http://dx.doi.org/10.1037/a0015558>

Huguet, P., Dumas, F., Monteil, J.M. and Genestoux, N. (2001) Social comparison choices in the classroom: Further evidence for students' upward comparison tendency and its beneficial impact on performance . *European Journal of Social Psychology* , 31, 557–578. <http://dx.doi.org/10.1002/ejsp.81>

Jones, S. and Regan, D. (1974) Ability evaluation through social comparison . *Journal of Experimental Social Psychology* , 10, 133–146. <http://dx.doi.org/10.1016/0022-1031%2874%2990062-6>

Kelley, H. (1967) Attribution theory in social psychology . In D. Levine (ed.) *Nebraska Symposium on Motivation* , pp. 192–238. Lincoln, NE: University of Nebraska Press.

Kelley, H. (1973) The process of causal attribution . *American Psychologist* , 28, 107–128. <http://dx.doi.org/10.1037/h0034225>

Klayman, J. and Ha, Y.W. (1987) Confirmation, disconfirmation and information in hypothesis-testing . *Psychological Review* , 94, 211–228. <http://dx.doi.org/10.1037/0033-295X.94.2.211>

Kruglanski, A. and Mayseless, O. (1990) Classic and current social comparison research: Expanding the perspective . *Psychological Bulletin* , 108, 195–208. <http://dx.doi.org/10.1037/0033-2909.108.2.195>

Kulik, J.A. and Mahler, H.I.M. (1987) Effects of preoperational roommate assignment on preoperative anxiety and postoperative recovery from coronary bypass surgery . *Health Psychology* , 6, 525–543. <http://dx.doi.org/10.1037/0278-6133.6.6.525>

Latané, B. (ed.) (1966) Studies in social comparison . *Journal of Experimental Social Psychology* , Suppl. 1.

Latané, B., Eckman, J. and Joy, V. (1966) Shared stress and interpersonal attraction . *Journal of Experimental Social Psychology* Suppl. 1 , 2, 92–102

Lockwood, P. and Kunda, Z. (1997) Superstars and me: Predicting the impact of role models on the self . *Journal of Personality and Social Psychology* , 73, 91–103. <http://dx.doi.org/10.1037/0022-3514.73.1.91>

Lombardi, W.J., Higgins, E.T. and Bargh, J.A. (1987) The role of consciousness in priming effects on categorization: Assimilation versus contrast as a function of awareness of the priming task . *Personality and Social Psychology Bulletin* , 13, 411–429. <http://dx.doi.org/10.1177/0146167287133009>

Major, B., Testa, M. and Blysm, W. (1991) Responses to upward and downward social comparisons: The impact of esteem-relevance and perceived control . In J. Suls, ed. and T.A. Wills (eds), *Social Comparison: Contemporary Theory and Research* , pp. 237–260. Hillsdale, NJ: Lawrence Erlbaum Associates.

Manis, M., Biernat, M. and Nelson, T.F. (1991) Comparison and expectancy processes in human judgment . *Journal of Personality and Social Psychology* , 61, 203–211. <http://dx.doi.org/10.1037/0022-3514.61.2.203>

Marsh, H.W. and Hau, K. (2003) Big-fish-little-pond-effect on academic self-concept. A cross-cultural (26 country) test of the negative effects of academically selective schools . *American Psychologist* , 58, 364–376. <http://dx.doi.org/10.1037/0003-066X.58.5.364>

Martin, R. (2000). ‘Can I do X?’ Using the proxy comparison model to predict performance. In J. Suls, ed. and L. Wheeler (eds), *Handbook of Social Comparison: Theory and Research* , pp. 67–80. New York: Kluwer/Plenum. http://dx.doi.org/10.1007/978-1-4615-4237-7_4

Martin, R., Suls, J. and Wheeler, L. (2002) Ability evaluation by proxy: The role of maximum performance and related attributes in social comparison . *Journal of Personality and Social Psychology* 82, 781–791. <http://dx.doi.org/10.1037/0022-3514.82.5.781>

Morse, S. and Gergen, K.J. (1970) Social comparison, self-consistency, and the concept of the self . *Journal of Personality and Social Psychology* , 16, 148–156. <http://dx.doi.org/10.1037/h0029862>

Mussweiler, T. (2003) Comparison processes in social judgment: Mechanisms and consequences . *Psychological Review* , 110, 472–489. <http://dx.doi.org/10.1037/0033-295X.110.3.472>

Mussweiler, T. and Bodenhausen, G. (2002) I know you are, but what am I? Self-evaluative consequences of judging in-group and out-group members . *Journal of Personality and Social Psychology* , 82, 19–32. <http://dx.doi.org/10.1037/0022-3514.82.1.19>

Mussweiler, T., Epstrude, K. and Ruter, K. (2005) The knife that cuts both ways: Comparison processes in social perception . In M. Alicke, ed. , D. Dunning, ed. , and J. Krueger (eds), *The Self in Social Judgment* , pp. 109–130. New York: Psychology Press.

Mussweiler, T. and Strack, F. (2000) The ‘Relative Self’: Informational and judgmental consequences of comparative self-evaluation . *Journal of Personality and Social Psychology* , 79, 23–38. <http://dx.doi.org/10.1037/0022-3514.79.1.23>

Neely, J.H. (1976) Semantic priming and retrieval from lexical memory: Evidence for facilitory and inhibitory processes . *Memory and Cognition* , 4, 648–654. <http://dx.doi.org/10.3758/BF03213230>

Neely, J.H. (1991) Semantic priming effects in visual word recognition: A selective review of current findings and theories . In D. Besner, ed. and G.W. Humphreys (eds), *Basic Processes in Reading: Visual Word Recognition* , pp. 264–336. Hillsdale, NJ: Lawrence Erlbaum.

Parducci, A. (1974) Contextual effects: A range-frequency analysis . In E.C. Carterette, ed. and M.P. Friedman (eds), *Handbook of Perception* , 7, 127–141. New York: Academic Press.

Pettigrew, T. (1967) Social evaluation theory: Convergences and applications . In D. Levine (ed.) *Nebraska Symposium on Motivation* , pp. 241–318. Lincoln, NE: University of Nebraska Press.

Radloff, R. (1966) Social comparison and ability evaluation . *Journal of Experimental Social Psychology*, Suppl. 1 , 2, 6–26. <http://dx.doi.org/10.1016/0022-1031%2866%2990061-8>

Schachter, S. (1959) *The Psychology of Affiliation* . Stanford, CA: Stanford University of Press.

Smith, W. and Sachs, P. (1997) Social comparison and task prediction: Ability similarity and the use of a proxy . *British Journal of Social Psychology* , 36, 587–602. <http://dx.doi.org/10.1111/j.2044-8309.1997.tb01151.x>

Spiegel, D., Bloom, J. and Yalom, I. (1981) Group support for patients with metastatic cancer: A randomized prospective outcome study . *Archives of General Psychiatry* , 38, 527–533. <http://dx.doi.org/10.1001/archpsyc.1980.01780300039004>

Srull, T.K. and Wyer, R.S. (1979) The role of category accessibility in the interpretation of information about persons: Some determinants and implications . *Journal of Personality and Social Psychology* , 37, 1660–1672. <http://dx.doi.org/10.1037/0022-3514.37.10.1660>

Stanton, A., Danoff-Burg, S., Cameron, C., Snider, P. and Kirk, S. (1999) Social comparison and adjustment to breast cancer: An experimental examination of upward affiliation and downward evaluation . *Health Psychology* , 18, 151–158. <http://dx.doi.org/10.1037/0278-6133.18.2.151>

Stapel, D.A. (2007) In the mind of the beholder: The Interpretation Comparison Model of accessibility effects . In D.A. Stapel, ed. and J. Suls (eds), *Assimilation and Contrast in Social Psychology* , pp. 143–164. New York: Psychology Press.

Stapel, D.A. and Koomen, W. (2000) Distinctiveness of others, mutability of selves: Their impact on self-evaluations . *Journal of Personality and Social Psychology* 79, 1068–1087. <http://dx.doi.org/10.1037/0022-3514.79.6.1068>

Stapel, D.A. and Koomen, W. (2001b) The impact of interpretation versus comparison goals on knowledge accessibility effects . *Journal of Experimental Social Psychology* 37, 134–149. <http://dx.doi.org/10.1006/jesp.2000.1440>

Stapel, D.A. and Suls, J. (2004) Method matters: Effects of explicit versus implicit social comparisons on activation, behavior, and self-views . *Journal of Personality and Social Psychology* , 87, 860–875. <http://dx.doi.org/10.1037/0022-3514.87.6.860>

Suls, J. (1986) Notes on the occasion of Social Comparison Theory's thirtieth birthday . *Personality and Social Psychology Bulletin* , 12, 289–296. <http://dx.doi.org/10.1177/0146167286123004>

Suls, J., Gastorf, J. and Lawhorn, J. (1978) Social comparison choices for evaluating a sex- and age-related ability . *Personality and Social Psychology Bulletin* , 4, 102–105. <http://dx.doi.org/10.1177/014616727800400121>

Suls, J., Martin, R. and Wheeler, L. (2000) Three kinds of opinion comparison: The Triadic Model . *Personality and Social Psychology Review* , 4, 219–237. http://dx.doi.org/10.1207/S15327957PSPR0403_2

Suls, J., Martin, R. and Wheeler, L. (2002) Social comparison: Why, with whom and with what effect? *Current Directions in Psychological Science* , 11, 159–163. <http://dx.doi.org/10.1111/1467-8721.00191>

Suls, J. and Miller, R.L. (1977) Social Comparison Processes: Theoretical and Empirical Perspectives . Washington, DC: Hemisphere.

Suls, J. and Miller, R.L. (1978) Ability comparison and its effects on affiliation preferences . *Human Relations* , 31, 267–282. <http://dx.doi.org/10.1177/001872677803100305>

Suls, J. and Tesch, F. (1978) Students' preferences for information about their test performance: A social comparison study . *Journal of Applied Social Psychology* , 8, 189–197. <http://dx.doi.org/10.1111/j.1559-1816.1978.tb00776.x>

Suls, J., ed. and Wheeler, L. (eds) (2000) Handbook of Social Comparison . New York: Kluwer/Plenum. <http://dx.doi.org/10.1007/978-1-4615-4237-7>

Suls, J. and Wheeler, L. (2007) Psychological magnetism: A brief history of assimilation and contrast in psychology . In D. Stapel, ed. and J. Suls (eds), Assimilation and Contrast in Social Psychology pp. 9–44. New York: Psychology Press.

Suls, J. and Wheeler, L. (2008) A reunion for approach/avoidance motivation and social comparison . In A. Elliot (ed.) *Handbook of Approach and Avoidance Motivation* . Mahwah, NJ: Lawrence Erlbaum Associates.

Taylor, S. and Lobel, M. (1989) Social comparison under threat: Downward evaluation and upward contacts . *Psychological Review* , 96, 569–575. <http://dx.doi.org/10.1037/0033-295X.96.4.569>

Tennen, H., McKee, T.E. and Affleck, G. (1991) Social comparison processes in health and illness . In J. Suls, ed. and L. Wheeler (eds), *Handbook of Social Comparison: Theory and Research* , pp. 443–486. New York: Kluwer Academic/Plenum.

Tesser, A. (1988) Toward a self-evaluation maintenance model of social behavior . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* 21, 181–227. New York: Academic Press.

Testa, M. and Major, B. (1990) The impact of social comparisons after failure: The moderating effects of perceived control . *Basic and Applied Social Psychology* , 11, 205–218. http://dx.doi.org/10.1207/s15324834basp1102_7

Turner, J. (1991) *Social Influence* . Belmont, CA: Brooks/Cole.

Wheeler, L. (1966) Motivation as a determinant of upward comparison . *Journal of Experimental Social Psychology*, Suppl. 1 , 2, 27–31. <http://dx.doi.org/10.1016/0022-1031%2866%2990062-X>

Wheeler, L. (2000) Individual differences in social comparison . In J. Suls, ed. and L. Wheeler (eds), *Handbook of Social Comparison* , pp. 141–158. New York: Kluwer Academic/Plenum. http://dx.doi.org/10.1007/978-1-4615-4237-7_8

Wheeler, L. and Koestner, R. (1984) Performance evaluation: On choosing to know the related attributes of others when we know their performance . *Journal of Experimental Social Psychology* , 20, 263–271. <http://dx.doi.org/10.1016/0022-1031%2884%2990051-9>

Wheeler, L., Koestner, R. and Driver, R. (1982) Related attributes in the choice of comparison others: It's there, but it isn't all there is . *Journal of Experimental Social Psychology* , 18, 489–500. <http://dx.doi.org/10.1016/0022-1031%2882%2990068-3>

Wheeler, L., Martin, R. and Suls, J. (1997) The proxy social comparison model for self-assessment of ability . *Personality and Social Psychology Review* , 1, 54–61. http://dx.doi.org/10.1207/s15327957pspr0101_4

Wheeler, L. and Miyake, K. (1992) Social comparison in everyday life . *Journal of Personality and Social Psychology* , 62, 760–733. <http://dx.doi.org/10.1037/0022-3514.62.5.760>

Wheeler, L. and Suls, J. (2005) Social comparison and self-evaluations of competence . In A. Elliot, ed. and C. Dweck (eds), *Handbook of Competence and Motivation* . New York: Guilford Press.

Wheeler, L. and Suls, J. (2007) Assimilation in social comparison: Can we agree on what it is? *Revue Internationale de Psychologie Sociale* , 20, 31–51.

Wills, T.A. (1981) Downward comparison principles in social psychology . *Psychological Bulletin* , 90, 245–271. <http://dx.doi.org/10.1037/0033-2909.90.2.245>

Wills, T.A. and Suls, J. (1991) Commentary: neo-social comparison and beyond . In J. Suls, ed. and T.A. Wills (eds), *Social Comparison: Contemporary Theory and Research* , pp. 395–412. Hillsdale, NJ: Lawrence Erlbaum.

Wood, J.V. (1989) Theory and research concerning social comparison of personal attributes . *Psychological Bulletin* , 106, 231–248. <http://dx.doi.org/10.1037/0033-2909.106.2.231>

Wood, J.V. (1996) What is social comparison and how should we study it? *Personality and Social Psychology Bulletin* , 22, 520–537. <http://dx.doi.org/10.1177/0146167296225009>

Wood, J.V., Taylor, S. and Lichtman, R. (1985) Social comparison in adjustment to breast cancer . *Journal of Personality and Social Psychology* , 49, 1169–1183. <http://dx.doi.org/10.1037/0022-3514.49.5.1169>

Zanna, M., Goethals, G.R. and HillS. (1975) Evaluating a sex related ability: Social comparison with similar others and standard setters . *Journal of Experimental Social Psychology* 11, 86–93. <http://dx.doi.org/10.1016/S0022-1031%2875%2980013-8>

<http://dx.doi.org/10.4135/9781446249215.n23>

Handbook of Theories of Social Psychology: Volume 1

Regulatory Focus Theory

Contributors: E. Tory Higgins

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Regulatory Focus Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n24>

Print pages: 483-505

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n24>

[p. 483 ↓]

Chapter 23: Regulatory Focus Theory

E. Tory Higgins, ed.

Abstract

Regulatory focus theory was the child of self-discrepancy theory and the parent of regulatory fit theory. As the child of self-discrepancy theory, it distinguishes between self-regulation in relation to hopes and aspirations (promotion ideals) versus self-regulation in relation to duties and obligations (prevention oughts). But in regulatory focus theory, promotion and prevention orientations are states that vary not only predispositionally across individuals but also can be situational induced. And the emotional and motivational effects of success and failure can be on a current task rather than just chronic congruencies and discrepancies to ideal and ought self-guides. Most importantly, the emphasis is on strategic differences between promotion and prevention in how desired states are attained, with promotion preferring eager means of advancement and prevention preferring vigilant means of maintenance. This emphasis on strategic differences also distinguishes regulatory focus theory from control system theories' concern with approach and avoidance at the system level rather than the strategic level. The asymmetry between promotion and prevention in strategic preferences and in the motivational effects of success and failure, with success strengthening motivation in promotion but failure strengthening motivation in prevention, gave birth to the regulatory fit idea that the manner of goal pursuit can sustain or disrupt a self-regulatory orientation.

Introduction

I moved from New York University (NYU) to Columbia University in 1989. The decision was very difficult, mostly because of the wonderful qualities of both options – an approach-approach conflict. The conflict would have been even worse if I had recognized and fully considered all of the consequences of each option. But I didn't. I had other things on my mind. My wife and I were about to be blessed with our first child and this was what I was thinking about. And, not surprisingly, our child was what I continued to think about for several months after she was born. It was not until my second term at Columbia (Spring, 1990) that I finally began thinking about what research I would do at Columbia. And when I did begin thinking about it, I was thunderstruck. I suddenly [p. 484 ↓] realized that I could no longer do the kind of research that I had done at NYU for years. I could no longer work on self-discrepancy theory. Like an irresponsible parent, I had thoughtlessly abandoned my theory by moving to Columbia. What was I going to do?

What struck me at that painful moment was that the research conditions at Columbia were completely different from those at NYU. At NYU, where I had developed self-discrepancy theory, there was a very large subject pool – thousands of potential participants per year. In contrast, the subject pool at Columbia was just a few hundred. And there was another difference that was at least as important. At NYU all students in the subject pool filled out a battery booklet that contained not only demographic questions but also personality questionnaires. You guessed it – one of those questionnaires was the Selves Questionnaire that my lab used to measure individuals' self-discrepancies. This meant that even before my lab contacted potential participants for a study, we already had information about their self-discrepancies. We could select participants as a function of whether they were high or low in different kinds of self-discrepancies. At Columbia, on the other hand, there was no such battery booklet. My lab would have to do a formal study with hundreds of participants just to obtain the background information on their self-discrepancies. By the time we had done that, we would have no more subject hours for research. It was hopeless!

So there I was, mourning over the lost of my theory and wondering what kind of research I would do instead. My first thought was, "Well, I will just have to go back

to doing research on priming and accessibility full-time.” It was a comforting thought, but not comforting enough. I still missed self-discrepancy theory. After all, whereas construct accessibility theory was an independent teenager by this time (Higgins et al., 1977; see [Chapter 4](#), this volume), self-discrepancy theory was still only a child (Higgins, 1987).

It was during this period, under these trying circumstances, that I began to find a solution to the problem of what to do about self-discrepancy theory – a solution that ultimately became regulatory focus theory. Without question, it was trying to continue parenting self-discrepancy theory that gave birth to regulatory focus theory (see Higgins, 2004, 2006a). I realized that what I needed was to find a way to test self-discrepancy theory without having to measure people's chronic self-discrepancies. In thinking about this problem, and ultimately finding a solution, regulatory focus theory was born, although I didn't appreciate that immediately. It took me some time to recognize what had happened (a separate theory construction issue I will discuss later). But now, I need to begin at the beginning. For you to appreciate the dilemma I faced in the spring of 1990, I need to provide some basic background information about self-discrepancy theory and describe the kinds of studies we had conducted to test it.

Self-Discrepancy Theory: The Parent of Regulatory Focus Theory

Why do people emotionally react so differently to the same tragic event? More specifically, why is it that when people are emotionally overwhelmed by a serious setback in their life, such as the death of their child, the loss of their job, or the breakup of their marriage, some suffer from depression whereas others suffer from anxiety? Self-discrepancy theory was developed to answer this question. Self-discrepancy theory proposed that even when people have the same specific goals, such as seniors in high school wanting to go to a good college or older adults wanting a good marriage, they often vary in how they represent these goals. The goals or standards that direct or guide our self-regulation are called *self-guides* in self-discrepancy theory. Some of us represent our self-guides as hopes or aspirations, the kind of person we ideally want

to be – *ideal* [p. 485 ↓] self-guides. Others of us represent our self-guides as duties or obligations, the kind of person we believe we ought to be – *ought* self-guides.

According to self-discrepancy theory (Higgins, 1987), it is the difference between failing to meet our ideals versus failing to meet our oughts that provides the key to unlocking the mystery of why we have different emotional reactions to the same negative life event. Self-discrepancy theory proposes that when a negative life event happens to us, it is represented as saying something about how we are doing. We compare our current, actual self to our self-guides: “Compared to the kind of person I want to be (e.g., going to a good college; having a good marriage), how am I doing?” We suffer emotionally when there is a discrepancy between our actual self and a self-guide – a *self-discrepancy*. When our actual self is discrepant from an ideal self-guide, we feel sad, disappointed, discouraged – dejection-related emotions that relate clinically to depression. When our actual self is discrepant from an ought self-guide, we feel nervous, tense, and worried – agitation-related emotions that relate clinically to anxiety disorders. According to self-discrepancy theory, then, our vulnerabilities to different kinds of emotional suffering depend on which type of self-guide is emphasized in our self-regulation – dejection/depression suffering when ideals are emphasized, and agitation/anxiety suffering when oughts are emphasized.

Research with clinically depressed and clinically anxious patients has found support for these proposals about emotional vulnerabilities. Discrepancies between patients' actual selves and their ideal self-guides predicts their suffering from depression more than it predicts their suffering from anxiety disorders, whereas discrepancies between patients' actual selves and their ought self-guides predicts their suffering from anxiety disorders more than it predicts their suffering from depression (Strauman, 1989). Because some individuals have actual-self discrepancies from both their ideal and ought self-guides, they can suffer from both depression and anxiety disorders.

At any one time, however, either individuals' ideal self-guides or their ought self-guides can be more accessible, and whichever is more accessible will determine which emotional syndrome they experience. This means that momentary situations can determine which syndrome is experienced by priming or activating either ideal or ought self-guides. For example, there is evidence that either actual-ideal discrepancies or actual-ought discrepancies can be made temporarily more accessible by exposing

individuals either to words which relate to an ideal they possess or to an ought they possess. When such priming of either an ideal or an ought occurs in an experiment, participants whose actual-ideal discrepancy is activated suddenly feel sad and disappointed and fall into a depression-related state of low activity (e.g., talk slower). In contrast, participants whose actual-ought discrepancy is activated suddenly feel nervous and worried and fall into an anxiety-related state of high activity (e.g., talk quicker). And these kinds of effects have been found with both clinical samples (Strauman, 1989) and non-clinical samples (Strauman and Higgins, 1987).

What is the psychological mechanism that underlies these effects? Self-discrepancy theory proposes that different emotions relate to different psychological situations that people experience. That is, the psychological situations produced by success or failure to meet our *ideals* are different from the psychological situations produced by success or failure to meet our *oughts*. Specifically, when events are related to our ideal self-guides (i.e., to our hopes and aspirations), we experience success as the presence of a positive outcome (a gain), which is a happy experience, and we experience failure as the absence of positive outcome (a nongain), which is a sad experience. In contrast, when events are related to our ought self-guides (i.e., our beliefs about our duties and obligations), we experience success as the absence of [p. 486 ↓] a negative outcome (a nonloss), which is a relaxing experience, and we experience failure as the presence of a negative outcome (a loss), which is a worrying experience. Consistent with this proposed underlying mechanism, research has shown (e.g., Higgins and Tykocinski, 1992) that individuals with strong ideals remember better events that reflect the absence or the presence of positive outcomes (gains and nongains), whereas individuals with strong oughts remember better events that reflect the presence or absence of negative outcomes (nonlosses and losses). People also remember better those events in their own lives that relate to whichever type of self-guide is more accessible for them (Strauman, 1992).

What kind of parenting is likely to result in children having either strong ideal self-guides or strong ought self-guides? In answering these questions, self-discrepancy theory relies on the basic idea that self-regulation in relation to ideal self-guides versus ought self-guides involves experiencing different psychological situations. When children interact with their parents (or other caretakers), the parents respond to their child in ways that make the child experience one of the different kinds of psychological

situations. Over time, the children respond to themselves like their parents respond to them, producing the same specific kinds of psychological situations, and this develops into the kind of self-guide (ideal or ought) that is associated with those psychological situations (see Higgins, 1991).

What pattern of parenting, then, predicts the development of strong ideal self-guides in children? It is when parents combine bolstering (when managing success) and love withdrawal (when disciplining failure). Bolstering occurs, for instance, when parents encourage the child to overcome difficulties, hug and kiss them when they succeed, or set up opportunities for the child to engage in success activities – it creates an experience of the presence of positive outcomes in the child. Love withdrawal occurs, for instance, when parents end a meal when the child throws some food, take away a toy when the child refuses to share it, stop a story when the child is not paying attention – it creates an experience of the absence of positive outcomes in the child.

What pattern of parenting predicts the development of strong ought self-guides in children? It is when parents combine prudence (when managing success) and punitive/critical (when disciplining failure). Prudence occurs, for instance, when parents “child-proof” the house, train the child to be alert to potential dangers, or teach the child to “mind your manners” – it creates an experience of the absence of negative outcomes in the child. Punitive/critical occurs, for instance, when parents play roughly with the child to get his or her attention, yell at the child when he or she doesn't listen, criticize the child when he or she makes a mistake – it creates an experience of the presence of negative outcomes. Indeed, consistent with these self-discrepancy theory predictions, there is recent evidence of positive associations between the critical and punitive parenting style and prevention-focused self-regulation and between the bolstering parenting style and promotion-focused self-regulation (see Higgins and Silberman, 1998; Keller, 2008; Manian et al., 2006).

In addition to distinguishing between ideal and ought self-guides, self-discrepancy theory distinguishes between different standpoints that can be taken in self-regulation (Higgins, 1987) – between self-regulation from our own independent standpoint (“What are my own goals and standards for myself?”) and self-regulation from the standpoint of a *significant other* person in our lives (e.g., “What are my mother's goals and standards for me?”). For example, in North America at least, there is evidence

that discrepancies from independent self-guides are a more important determinant of emotional vulnerabilities for males than for females. In contrast, discrepancies from significant other self-guides are more important for females than for males (Moretti and Higgins, 1999a, 1999b).

[p. 487 ↓]

Now Back to the Story: The Birth of Regulatory Focus Theory

Research testing self-discrepancy theory had always used the participants' self-discrepancies as part of the study. Most studies had tested the effects of self-discrepancies on emotional responses – the relation between self and affect – but occasionally we took advantage of previous work on priming and accessibility to make either actual-ideal or actual-ought discrepancies temporarily more accessible by priming either ideal or ought self-guides, respectively. For example, in one study by Higgins et al. (1986) participants were selected who were either high in *both* actual-ideal and actual-ought discrepancies or low in both, and when they arrived for the study they were asked either to discuss their own and their parents' hopes and aspirations for them (ideal priming) or to discuss their own and their parents' beliefs concerning their duties and obligations (ought priming). The study found that for "high in both" participants, but not "low in both" participants, ideal priming produced dejection-related emotions whereas ought priming produced agitation-related emotions.

The idea of using priming to make temporarily more accessible either ideal or ought self-guides (i.e., ideal or ought goals) was itself new to us. In the beginning, our research was based solely on using the Selves Questionnaire measure of self-discrepancies, which ideographically measured individuals' stable self-discrepancies (e.g., Higgins et al., 1985). It took a while for us to realize that we could take advantage of what we knew about accessibility and priming (see Chapter 4, this volume) to activate different kinds of self-discrepancies, thereby having more experimental control over our tests of self-discrepancy theory by activating momentarily either actual-ideal or actual-

ought discrepancies. This experimental method discovery turned out to be a critical turning point for the birth of regulatory focus theory, but we did not know that at the time.

At the time, we still thought of self-discrepancy theory in strictly personality terms. We were using priming to make one kind of chronic self-discrepancy more accessible than another; that is, make *it* the more currently active self-discrepancy. We knew from research that was going on during the same period that temporary accessibility from situational priming could trump, at least for a while, chronic accessibility from established individual differences (Bargh et al., 1988). But self-discrepancy theory was still about chronic individual differences in stored self-discrepancies. For individuals with both ideal and ought discrepancies, as in the Higgins et al. (1986) study, we were simply using the priming to make one or the other discrepancy more active at the moment.

So here I was in 1990 realizing that I could not conduct studies at Columbia which required measuring participants' self-discrepancies beforehand. To me that meant I could no longer do research on self-discrepancy theory. What was I to do? As I mentioned earlier, one thought was to return to priming and accessibility research full-time. But I did not want to abandon self-discrepancy theory so early in its development. Perhaps I was thinking about priming and accessibility while wanting to continue working on ideals and oughts that was critical. I don't know. What I do know is that I suddenly realized that there was a solution to my problem. What had fascinated me about accessibility for a long time was the fact that accessibility was a state, and individuals did not know the source of their accessibility state – it could be from chronic accessibility, priming, or both (see [Chapter 4](#), this volume). I had thought about accessibility being a *common language* for variability across persons (chronic accessibility) and variability across situations (priming), which provided a different perspective on the classic "person-situation" debate (Higgins, 1990). I had also thought that standards provided another common language for variability across persons (personal standards) and variability across situations (contextual standards) [[p. 488 ↓](#)] (Higgins, 1990). But I had *not* fully appreciated what the notion of a common language implied nor had I thought about how accessibility and standards could be combined.

I now realized that the distinction between ideal and ought self-regulation was being unnecessarily restricted by self-discrepancy theory. Rather than this distinction being

about chronic discrepancies (or congruencies) between the actual self and ideal or ought self-guides, it was more generally about *two different systems of self-regulation*. At any moment, people could be in a state of regulating in relation to hopes or wishes (ideals) or they could be in a state of regulating in relation to duties or responsibilities (oughts). And, importantly, this could be true regardless of whether they did or did not possess chronic ideal or ought discrepancies.

From this broader, two-distinct-systems perspective, the emotional and motivational implications of individuals being in a *state* of ideal self-regulation versus a *state* of ought self-regulation could be studied *without any need to measure individuals' self-discrepancies*. What mattered now was the distinction between two systems of self-regulation, and one or the other *system* could be activated through priming or other experimental manipulations. It was *not* necessary to measure self-discrepancies in order to study the implications of ideal versus ought self-regulation. My problem was solved!

An issue did remain, however. Specifically, would I still be testing self-discrepancy theory – that is, *the theory itself* – if I was no longer measuring self-discrepancies? For many years I have taught a course on theory construction in psychology (see Higgins, 2004). I was well aware in 1990 that it was important to distinguish between an *extension* of a theory versus a *new* theory. Failure to do so can be very unfair to both the “old” theory and the “new” theory. The question, then, was whether there were two theories – the old self-discrepancy theory and the new “X” theory – or simply earlier and later versions of self-discrepancy theory. Why treat what might be just an extension or elaboration of the extant self-discrepancy theory as a new theory? I have noted elsewhere that this is a gray area (Higgins, 2004), and it comes down to whether one believes that the “new” theory, in fact, really adds something that is fundamentally new.

In brief, I did believe that something fundamentally new had been added by the “new,” and as-yet-unnamed, theory. The “new” theory – what became called *regulatory focus theory* – was concerned with distinct self-regulatory *states* that varied across *both persons and situations*. Unlike self-discrepancy theory, it was *not* a personality theory. Whereas ideal and ought self-guides varied chronically across persons in self-discrepancy theory, ideal and ought self-regulation in regulatory focus theory varied across situations as well as persons. And this difference really mattered. It inspired

research that would not have been generated by self-discrepancy theory, such as framing effects on problem solving and decision making (see Higgins, 1998).

I believed then, as I believe now, that what distinguishes a theory are the discoveries that it generates (Higgins, 2004). And it was clear even then that this new theory would generate studies that could make discoveries that self-discrepancy theory would not. What made the most sense, and was most fair, was to let self-discrepancy theory continue to develop in its same basic form – after all it was still just a child – while treating its offspring as something deserving its own separate path of development. And self-discrepancy theory did continue to develop (see Moretti and Higgins, 1999a, 1999b), such as considering the overlap between own standpoint and significant other standpoint on someone's goals in order to distinguish among independent self-regulation (just own standpoint), identified self-regulation (shared reality from own and significant other overlap), and introjected self-regulation (just significant other "felt presence of other"). Moreover, when regulatory focus theory began to mature, it was used to enrich self-discrepancy theory and broaden its implications, as evident in [p. 489 ↓] Tim Strauman's construction of *self-system therapy* (Strauman et al., 2006). Like all good children, then, regulatory focus theory had a positive impact on its parent. Parent and child have each benefited enormously from one another – a nice family story. It is time, then, to describe what this new regulatory focus theory was all about and discuss the new research directions that it inspired.

Promotion and Prevention Systems of Self-Regulation

From the ancient Greeks, through seventeenth- and eighteenth-century British philosophers, to twentieth-century psychologists (see Kahneman et al., 1999), the hedonic principle that people are motivated to approach pleasure and avoid pain has dominated our understanding of people's motivation. It is the basic motivational assumption of theories across all areas of psychology, including theories of emotion in psychobiology (e.g., Gray, 1982), conditioning in animal learning (e.g., Mowrer, 1960; Thorndike, 1935), decision-making in cognitive psychology (e.g., Edwards, 1955; Kahneman and Tversky, 1979), consistency in social psychology (e.g., Festinger,

1957; Heider, 1958), and achievement motivation in personality (e.g., Atkinson, 1964). Even when Freud, in *Beyond the Pleasure Principle* (Freud, 1950/1920), assigned a motivational role to the ego's reality principle, he made it clear that the reality principle "at bottom also seeks pleasure although a delayed and diminished pleasure" (Freud, 1952/1920: 365). Perhaps the clearest statement on the importance of hedonic experiences to motivation was given by Jeremy Bentham (1781/1988: 1): "Nature has placed mankind under the governance of two sovereign masters, *pain* and *pleasure*. It is for them alone to point out what we ought to do, as well as to determine what we shall do. On the one hand the standard of right and wrong, on the other the chain of causes and effects, are fastened to their throne."

Within this historical context, the contribution of self-discrepancy theory and, especially, regulatory focus theory has been to emphasize the significance of the different motivational systems that underlie pleasure and pain. It is not enough to know that people approach pleasure and avoid pain. It is critical to know *how* they do so. The starting assumption of regulatory focus theory is that the *different ways* that people approach pleasure and avoid pain can be *more* significant for understanding motivation and emotion than the hedonic principle per se.

To illustrate, for Gray (1982) and Mowrer (1960), as well as Carver and Scheier (1981, 1990 a, b), the important motivational distinction was between the approach system (BAS) and the avoidance or inhibition system (BIS). Gray (1982) and Mowrer (1960) explicitly included both approaching reward (the presence of a positive outcome) and approaching safety (the absence of a negative outcome) as *equivalent* cases of approaching a desired end-state. In contrast, self-discrepancy theory (Higgins, 1987) and regulatory focus theory (Higgins, 1996) explicitly distinguished between the promotion-ideal system versus the prevention-ought system as *two different ways of approaching desired end-states*. Within approach, and within avoidance, there were *distinct systems of self-regulation*, and, according to self-discrepancy theory and regulatory focus theory, it was the difference between the nature of these systems that mattered emotionally and motivationally rather than the hedonic principle per se.

The Significance of Regulatory Focus Theory for Emotional Experiences

As I discussed earlier, self-discrepancy theory distinguished between different types of pleasures and different types of pains as a function of which self-guide was involved in an actual self-congruency or discrepancy, [p. 490 ↓] cheerful-related and dejected-related feelings for ideal congruencies and discrepancies, and quiescent-related and agitated-related feelings for ought congruencies and discrepancies. Regulatory focus theory moved beyond this personality-based distinction by considering momentary successes and failures for people who happened to be in either a promotion focus state or a prevention focus state.

According to regulatory focus theory, the distinct emotional experiences from having a promotion focus versus a prevention focus are *not* restricted to personality. Anyone at a particular moment can be pursuing a goal with a promotion focus or a prevention focus. If a person in a promotion state succeeds in their goal pursuit, they will experience cheerful-related feelings (e.g., feel happy), and if they fail they will experience dejected-related feelings (e.g., feel sad). If a person in a prevention state succeeds, they will experience quiescent-related feelings (e.g., feel calm), and if they fail they will experience agitated-related feelings (e.g., feel tense) (see Idson et al., 2000).

The emotional significance of the two distinct self-regulatory systems was no longer restricted to actual-self relations to ideal and ought self-guides. It extended to any case of success or failure when self-regulating with a promotion focus or a prevention focus. One intriguing study which illustrates this included participants who did or did not believe that their father would strongly hope they would do well on the current task (promotion father) and other participants who did or did not believe that their father would view it as their obligation to do well on the task (prevention father). For “promotion father”-primed participants, success or failure on the current task produced feelings along the cheerful-dejected dimension, whereas for “prevention father”-primed participants success or failure produced feelings along the quiescent-agitated dimension (see Shah, 2003).

Although priming or activating a personal ideal or a personal ought is one way to induce a promotion or prevention state, it is not the only way. Other momentary situations can induce promotion or prevention states. In an early study by Roney et al. (1995), for example, success and failure feedback was framed in a promotion or prevention manner. In the promotion-framing condition, success and failure feedback was “you got that one” and “you didn’t get that one,” respectively. In the prevention-framing condition, success and failure feedback was “you didn’t miss that one” and “you missed that one,” respectively. By the end of the task, everyone had failed. As predicted by regulatory focus theory, agitation-related emotions increased from pretest to post-test more for prevention-framed participants than for promotion-framed participants, whereas the opposite was true for dejection-related emotions.

Even when individuals' promotion focus and prevention focus are chronic predispositions, it does not require that there be chronic self-discrepancies or self-congruencies in order for distinct emotions to be produced. Distinct emotions will be produced by success or failure on a current task. Our discovery of this fact involves another story with a message about theory construction and testing. It is remarkable, to me at least, how our knowledge in one scientific area is not readily applied to theory construction and theory testing in another scientific area. To illustrate, James Shah and I wanted to find a way to measure the chronic strength of individuals' promotion focus and prevention focus independent of whether they had self-discrepancies or self-congruencies. We felt that this was important because it could predict how they would feel from success or failure in a momentary situation independent of their past history of success and failure. We believed, for example, that whether individuals had actual-ideal discrepancies or not, if they pursued a goal with a promotion focus they would feel sad after failure and happy after success; and whether they had actual-ought discrepancies or not, if they pursued a goal with a prevention focus they [p. 491 ↓] would feel nervous after failure and relaxed after success.

For many months we tried various ways to measure the chronic strength of individuals' promotion focus and prevention strength that would predict the likelihood that they would pursue a goal in a promotion state or a prevention state. Time after time we developed what we thought was a very clever measure, but time after time the measure failed to work consistently. Then in 1993 I attended, as usual, the Social Psychology Winter Conference in Utah where I heard John Bassili describing his recent work on

attitude strength, work that was inspired by the earlier work of Russ Fazio (see Bassili, 1996; Fazio, 1986). This work took advantage of the notion of chronic accessibility.

By the time of that conference, I had worked on chronic accessibility for many years (e.g., Higgins et al., 1982; see [Chapter 4](#), this volume). But somehow it had never occurred to me to take advantage of what I knew about chronic accessibility to devise a measure of promotion and prevention strength. But now, after returning from the conference, James Shah and I did exactly that. In brief, we used individuals' response latencies when describing their personal ideals and oughts as a measure of the chronic accessibility of their ideals and oughts. Our assumption about ideals and oughts, like that of Bassili and Fazio about attitudes, was that chronic accessibility related positively to self-regulatory strength. Specifically, the more that a person's ideals had higher chronic accessibility than their oughts, the more that person's promotion system predominated; and the more that a person's oughts had higher chronic accessibility than their ideals, the more that person's prevention system predominated.

This implicit measure of promotion and prevention strength worked out very well. It turned out that promotion and prevention strength were independent of the extent to which individuals had ideal or ought discrepancies. Moreover, promotion and prevention strength moderated self-discrepancy effects, such that an actual-ideal discrepancy predicted suffering from dejected-related emotions more when promotion strength was also high, and an actual-ought discrepancy predicted suffering from agitated-related emotions more when prevention strength was also high (Higgins et al., 1997). This was a very nice parent and child collaboration that benefited both family members. Equally nice was now we had what we always wanted – an implicit measure of the likelihood that a person would pursue a goal with either a promotion focus (predominant promotion strength) or a prevention focus (predominant prevention strength), independent of the extent to which that person had actual-ideal or actual-ought discrepancies.

Individuals with predominant promotion strength – as measured by greater chronic accessibility of ideals than oughts – should emotionally experience success and failure along the cheerful-dejected dimension, and should emotionally appraise objects in the world along the cheerful-dejected dimension. In contrast, individuals with predominant prevention focus – as measured by greater chronic accessibility of oughts than ideals

– should emotionally experience success and failure along the quiescent-agitated dimension, and should emotionally appraise objects in the world along the quiescent-agitated dimension. All of these predictions were supported (see Idson et al., 2000; Shah and Higgins, 2001).

The Significance of Regulatory Focus Theory for Strategic Goal Pursuit

The bread and butter of self-discrepancy theory was to distinguish among different kinds of emotional experiences – different kinds of pleasure and different kinds of pain – and predict when they would occur. But even here, as I described above, regulatory focus theory had something new and important to add by predicting emotional [p. 492 ↓] effects of success and failure on a current task that varied as a function of situationally induced promotion or prevention states, and by predicting emotional effects of chronic promotion or prevention strength independent of chronic self-discrepancies. The major difference between self-discrepancy theory and regulatory focus theory, however, was regulatory focus theory's concern with understanding how the two distinct systems of self-regulation worked; how exactly the promotion and prevention systems *worked differently* to approach pleasure and avoid pain. Addressing this *how* issue would really move us beyond the hedonic principle and beyond self-discrepancy theory. This issue became the major concern of our lab in the early 1990s.

At the time that we began working on regulatory focus theory, Carver and Scheier had also developed a self-regulatory model that was concerned with motivation and emotion (Carver and Scheier, 1990a, 1990b). In their model there were also two self-regulatory systems – an *approach* system of reducing discrepancies to desired states as reference points, and an *avoidance* system of amplifying discrepancies from undesired states as reference points. As I mentioned earlier, this general distinction between an approach system and an avoidance system was a classic and important way of thinking about self-regulatory systems, and it was elaborated and developed by Carver and Scheier in significant and innovative directions (see Carver and Scheier, [Chapter 24](#), this handbook). What we needed to do conceptually and empirically was to clarify and demonstrate how the distinct systems identified by regulatory focus theory were *not* the

same as the distinction between an approach system versus an avoidance system as described in Carver and Scheier's model or in other approach-avoidance models, such as Mowrer's (1960) model, Gray's (1982) model, Atkinson's (1964) model, Lopes' (1987) model, and so on.

Carver and Scheier's distinction between an approach system and an avoidance system involves a distinction between two different reference points – a desired end-state as reference point versus an undesired end-state as reference point. I have called this a *regulatory reference* distinction (Higgins, 1997). (The distinction of Atkinson and Lopes between hope and fear is a *regulatory anticipation* distinction – see Higgins [1997].) In contrast, the regulatory focus theory distinction between a promotion focus system and a prevention focus system involves a distinction *within* a desired end-state as reference point and *within* an undesired end-state as reference point. Within a desired end-state as reference point, for example, individuals can have a promotion focus on ideals and accomplishments as the desired end-state or they can have a prevention focus on oughts and safety as the desired end-state. Regulatory focus and regulatory reference are orthogonal distinctions.

Continuing with a desired end-state as the reference point, how does self-regulation with a promotion focus differ from self-regulation with a prevention focus? Early on we decided that the difference between promotion versus prevention concerns would translate into a difference in *strategic* preferences for how to pursue goals. With a promotion focus, individuals would prefer the strategy of approaching self-states that were matches to a desired end-state. With a prevention focus, individuals would prefer the strategy of avoiding self-states that were mismatches to a desired end-state. This was a distinction at the *strategic* level instead of Carver and Scheier's distinction at the level of approach versus avoidance systems. *Within* Carver and Scheier's *approach* system, this was a distinction between a promotion strategy of approaching matches versus a prevention strategy of avoiding mismatches. *Within* Carver and Scheier's *avoidance* system, regulatory focus theory distinguished between a promotion strategy of approaching mismatches to an undesired end-state versus a prevention system of avoiding matches to an undesired end-state.

To test this new strategic distinction, an early study had participants read about [p. 493 ↓] many different episodes in the life of an individual that occurred over several days

(Higgins et al., 1994). In each of the episodes, the target was trying either to experience a desired end-state or not experience an undesired end-state. The episodes were of the following kinds:

- 1 *Approaching* matches to a *desired end-state*: “Because I wanted to be at school for the beginning of my 8:30 psychology class which is usually excellent, I woke up early this morning.”
- 2 *Avoiding* mismatches to a *desired end-state*: “I wanted to take a class in photography at the community center, so I didn’t register for a class in Spanish that was scheduled at the same time.”
- 3 *Approaching* mismatches to an *undesired end-state*: “I dislike eating in crowded places, so at noon I picked up a sandwich from a local deli and ate outside.”
- 4 *Avoiding* matches to an *undesired end-state*: “I didn’t want to feel tired during my very long morning of classes, so I skipped the most strenuous part of my morning workout.”

Using an “unrelated studies paradigm,” the participants first described either their personal ideals or their personal oughts to experimentally induce either a promotion focus or a prevention focus. Then they read the story and, afterward, tried to remember it. Across the desired and undesired end-states as reference points, the promotion-focused participants better remembered the episodes that involved strategic approach than the episodes that involved strategic avoidance, whereas the opposite was true for the prevention-focused participants.

Another study by Higgins et al. (1994) had participants make their own strategic choices regarding the desired end-state of friendship. The first phase of the study identified different friendship tactics. There were three tactics for the strategy of *approaching* matches: (a) “Be generous and willing to give of yourself”; (b) “Be supportive to your friends. Be emotionally supportive”; and (c) “Be loving and attentive.” There were also three tactics for the strategy of *avoiding* mismatches: (a) “Stay in touch. Don’t lose contact with friends”; (b) “Try to make time for your friends and not neglect them”; and (c) “Keep the secrets friends have told you and don’t gossip about friends.” In a later phase of the study, new participants varying in chronic strength of regulatory focus were

given all six tactics and were asked the *same* general question about friendship: “When you think about strategies for *friendship*, which THREE of the following strategies would you choose?” Across all participants, strategic approach tactics were chosen more than strategic avoidance tactics. But the study also found that predominant promotion participants chose more strategic approach tactics than predominant prevention participants, and predominant prevention participants chose more strategic avoidance tactics than predominant promotion participants.

The results of these two initial studies conducted in the early 1990s provided the first support for the unique prediction of regulatory focus theory that the promotion and prevention systems differed in the *strategic* manner of their goal pursuit. The distinction between the promotion and prevention systems was *not* a distinction between approach versus avoidance systems like Carver and Scheier's model and other earlier models. *Within* the approach system, and *within* the avoidance system, promotion and prevention differed in *how* goals were pursued. This strategic difference between the two regulatory focus systems was a critical distinguishing feature. It would have implications for self-regulation that were not fully foreseen at the time that these studies were done – implications that later led to the birth of regulatory fit theory (Higgins, 2000). I believe that the implications were not foreseen in part because the *terminology* of regulatory focus theory at that point in its development was more of a hindrance than a help. It is time to turn to that part of the story of regulatory focus theory.

[p. 494 ↓]

How “Promotion” and “Prevention” Became the Labels for the Two Systems

I have already talked about how theory development is a family affair in the sense of one theory giving birth to another, as self-discrepancy theory gave birth to regulatory focus theory. There are other ways in which theory development is a family affair (see Higgins, 2006a). For example, friends and colleagues provide feedback and suggestions at key stages in the development of a theory that have different kinds

of positive effects in its development. An illustration of this concerns the labeling of “regulatory focus theory” and its “promotion” and “prevention” systems. These were *not* the original labels.

In the summer of 1993, I attended a conference at Ringberg Castle that was organized by Peter Gollwitzer and John Bargh – a conference with the now classic titles, “For Whom the Ring Bergs” and “Four Days at Ringberg: Four Days at Ringberg” (see Gollwitzer and Bargh, 1996). At that conference, I discussed my new theory, called “regulatory outcome focus” that concerned two distinct self-regulatory systems, called “positive outcome focus” and “negative outcome focus” (Higgins, 1996; see also Roney et al., 1995). A couple of years later (1995), I attended another conference in Italy that was organized by Arie Kruglanski and Augusto Palmonari. I presented a similar talk to the one that I had presented at the Ringberg Conference. Sitting in the audience was Marilynn Brewer.

Marilynn came up to me after my talk and asked me why I was using the labels “positive outcome focus” and “negative outcome focus” in my new theory. As I remember our conversation, she said something like the following:

Don't those labels contradict the very point you are trying to make?
Aren't you arguing that both of your systems have pain and pleasure,
both have positive and negative outcomes? Isn't your point *not* to
confuse *your* two distinct systems with approaching positive outcomes
versus avoiding negative outcomes? But your labels, “positive outcome
focus” and “negative outcome focus,” make it sound like the distinction
is about positive versus negative outcomes.

It was hard to argue her point. Actually, a good life lesson with Marilynn is to listen carefully to what she has to say because you will learn something useful. When I arrived home I was convinced that I had to find new labels. But which labels? I knew that, developmentally, having a “positive outcome focus” was about “bolstering” and *nurturing advancements*. It was about fulfilling wishes and aspirations. For desired end-states, it was about strategically *approaching matches*. It did not take too long – with my trusty tool, *Roget's Thesaurus* to help me – to find a label that worked: *promotion*. That definitely sounded right; it had all the right connotations.

Now I needed a new label for “negative outcome focus.” I realized at this point that to follow Marilynn’s wise counsel I needed a label that not only captured the psychology of a “negative outcome focus” but also sounded positive. I needed two labels that were both positive and, at the system level, both involved approaching a desired end-state. It would be nice if the labels also implied a difference at the strategic level, but it was critical that both have positive valence at the system level.

With 20/20 hindsight, the answer may seem obvious to you now, but it was *not* obvious to me then. This was in part because, thanks to the self-discrepancy theory parentage, our research emphasis regarding “negative outcome focus” was on *ought* self-guides (i.e., duties and obligations). When I began searching for a new label, it was *ought* self-regulation that I had in mind – not safety and security. But, eventually, I came across the concept of *prevention*. Not only did “prevention” capture the association with *security* concerns and strategically avoiding *mismatches* (for desired end-states), but it was also a three-syllable word with “p” as the initial consonant – just like “promotion”! [p. 495 ↓] It was perfection (another three-syllable word with “p” as the initial consonant).

The theory was now about the distinction between self-regulation with a promotion focus and self-regulation with a prevention focus. To keep it simple, the theory should simply be called “regulatory focus theory.” To this day, I should note, papers appear that refer to the theory as “self-regulatory focus theory” – perhaps a more accurate phrase but too many words. When possible, try to find labels for your concepts, like “ideal” and “ought” or “promotion” and “prevention,” that are simple and everyday terms. (Donald Campbell’s [1958] label, “entitativity,” famously broke this rule but prospered anyway.)

I have always believed that words matter, and that exploring the different meanings of a word, its different denotations and connotations, is a great tool for discovering the psychological underpinnings of the concept to which the word refers. Indeed, for this reason, there was a kind of positive externality, an unexpected benefit, of changing the labels in the theory from “positive outcome focus” and “negative outcome focus” to “promotion focus” and “prevention focus.” For the reasons that Marilynn Brewer suggested, the original labels were more confusing than helpful because they differed from each other only by the words “positive” and “negative.” But this valence distinction

is precisely *not* what the theory is about. In contrast, the difference between “promotion” and “prevention” *is* precisely what the theory is about.

The more I thought about the words “promotion” and “prevention” – once again with the help of my trusty *Roget's Thesaurus* plus the *Webster* and *Oxford English* dictionaries – the more the differences between them became apparent. Perhaps most important initially, it occurred to me that the way to fulfill a promotion focus was by being *eager* and enthusiastic, whereas the way to fulfill a prevention focus was by being *vigilant* and careful. Up to this point I had emphasized the strategic difference between approaching matches to desired ideals and avoiding mismatches to desired oughts. This strategic distinction, I believe, is still accurate, but it is less generative than thinking of the difference between being eager in the service of promotion and being vigilant in the service of prevention. Indeed, this new way of thinking ultimately gave birth to another new child – *regulatory fit theory* (Higgins, 2000).

Regulatory Focus Theory: The Parent of Regulatory Fit Theory

I suppose that it would have been possible to think about approaching matches to a desired end-state as being a strategy that sustained (versus disrupted) a positive outcome focus and that avoiding mismatches to a desired end-state as being a strategy that sustained a negative outcome focus. But surely it is easier to think of an eager strategy as sustaining promotion and a vigilant strategy as sustaining prevention. And in the late 1990s, after the label changes, we began to do studies where we varied which strategies people used during goal pursuit (e.g., Förster et al., 1998). In one study by Shah et al. (1998), for example, participants solved green-colored anagrams in order to gain points versus red-colored anagrams in order not to lose points. We found that promotion-focused participants performed better on the (eager) green anagrams than the (vigilant) red anagrams, but the opposite was true for prevention-focused participants. Another study by Shah et al. (1998) found that predominant promotion participants performed better on an anagram task when the suggested strategy was to try and find 90 percent or more of the words (eager) versus when the suggested strategy was to try not to miss more than 10 percent of the words (vigilant).

Over time we began to think of these different strategies as being eager versus vigilant strategies, and we began to label them as such. Importantly, we thought of [p. 496 ↓] strategic eagerness as an approach strategy and strategic vigilance as an avoidance strategy. Because this strategic difference was a kind of *approach versus avoidance* distinction, it became all the more important to distinguish regulatory focus theory from control theories like Carver and Scheier's. Whereas Carver and Scheier's approach-avoidance distinction was at the system level (i.e., approaching desired end-states versus avoiding undesired end-states), our approach-avoidance distinction was at the strategic level. The different conceptions of avoidance at the system level versus the strategic level provided a critical test between these theories.

At Carver and Scheier's system level of avoidance, moving away from an undesired end-state should *decrease* avoidance intensity over time (see Miller, 1959). At regulatory focus theory's strategic level of avoidance, moving toward a desired end-state with prevention-focused vigilant avoidance should *increase* motivational intensity over time according to a "goal looms larger effect" (Lewin, 1935; Miller, 1959). These different predictions were tested in several studies using alternative measures of change in motivational intensity over time (i.e., arm pressure; persistence) and the results supported the prediction of regulatory focus theory (see Förster et al. 1998, 2001).

During the same general period we were conducting our "goal looms larger effect" studies, we were conducting other studies which suggested that the combination of promotion plus eager and prevention plus vigilance involved some sort of compatibility that had its own motivational significance. Earlier, Lorraine Idson, Nira Liberman, and I (Idson et al., 2000) had conducted studies in which participants imagined buying a book and choosing between paying with cash or paying with a credit card, with the book's price being lower if you paid with cash (see Thaler, 1980). The participants reported either how good they would feel if they paid with cash (the positive "success" outcome) or how bad they would feel if they paid with a credit card (the negative "failure" outcome). These earlier studies found that the pleasant feelings from the positive "success" outcome were more intense for promotion than prevention, that is, feeling cheerful versus feeling quiescent, whereas the painful feelings from the negative "failure" outcome were more intense for prevention than promotion, that is, feeling agitated versus feeling dejected.

In addition to measuring how good or bad participants felt about the outcome, our new studies included separate measures of pleasure/pain intensity and strength of motivational force (see Idson et al., 2004). For example, in one study that induced either a promotion focus or a prevention focus by priming either ideals or oughts, *pleasure-pain intensity* was measured by asking the participants how pleasant the positive outcome would be or how painful the negative outcome would be; and *strength of motivational force* was measured by asking the participants how motivated they would be to make the positive outcome happen (in the positive outcome condition) or how motivated they would be to make the negative outcome not happen (in the negative outcome condition).

We found that pleasure/pain intensity and strength of motivational force *each* made significant *independent* contributions to the perceived value of the imagined outcome (i.e., its goodness/badness). We also found that for the positive success outcome, strength of motivational force was higher in promotion than prevention; but for the negative failure outcome, strength of motivational force was higher in prevention than promotion. What these studies discovered was that there is an *asymmetry* between promotion and prevention with respect to whether success or failure yields a stronger motivational force. Other studies conducted around the same time also found this asymmetry. For promotion, there was better performance under conditions of success than failure. For prevention, there was better performance under conditions of failure than success [p. 497 ↓] (Förster et al., 2001; Idson and Higgins, 2000). This asymmetry between promotion and prevention regarding the effects of success versus failure was another major feature distinguishing the two systems. It was not a feature that had been identified in models distinguishing approach versus avoidance systems. It was something new about the difference between the promotion and prevention systems, and I knew it was important.

The earlier performance studies had found that goal pursuit in a promotion focus yielded better performance when an eager than a vigilant strategy was used, and the opposite was true for goal pursuit in a prevention focus. Now there was this new asymmetry for the effects on motivational strength of success versus failure. What was going on? The key to the solution was to recognize that the effects for intensity of emotions that we had found earlier (Idson et al., 2000) reflected differences in motivational strength. That is, when promotion-focused individuals feel cheerful after

success their motivation is high, but when they feel dejected after failure their motivation is low. In contrast, when prevention-focused individuals feel quiescent after success their motivation is low, but when they feel agitated after failure their motivation is high. And these differences in motivational strength are related to the same eagerness and vigilance that were involved in the earlier performance studies. That is, when promotion-focused individuals feel cheerful after success they are eager (strong motivation), but when they feel dejected after failure they are *not* eager (weak motivation). In contrast, when prevention-focused individuals feel agitated after failure they are vigilant (strong motivation), but when they feel quiescent after success they are *not* vigilant (weak motivation).

What united all the findings of these studies in the late 1990s was that being eager sustains motivational strength for individuals in a promotion focus, whereas being vigilant sustains motivational strength for individuals in a prevention focus. This solution to the mystery of what was going on in our recent regulatory focus studies gave birth to regulatory fit theory. What became clear was that there was another self-regulatory principle which contributed to the effects we were finding – the principle of *regulatory fit*.

Within this same period, a serendipitous event occurred. In 1999 I learned that I would receive the APA's Award for Distinguished Scientific Contributions, which meant that I would be giving a talk in 2000 that would then appear as an article in the *American Psychologist*. I wasn't sure whether to give a talk about past research testing regulatory focus theory or to give a talk on something new. I preferred the latter. But what would a new talk be about? I was quite excited about the findings from our recent studies and the solution to the mystery of what might underlie them. So I decided that my speech and paper for the APA award would present the new principle of regulatory fit. But this meant that I had to develop the new theory in short order.

The major proposal of regulatory fit theory was that people experience regulatory fit when the *manner* of their engagement in an activity *sustains* (versus disrupts) their current regulatory orientation (Higgins, 2000). An eager manner sustains promotion and a vigilant manner sustains prevention. But regulatory fit was *not* the same as regulatory focus because it concerned the relation between *any* goal pursuit orientation and the strategic manner in which the goal is pursued. Indeed, when I reconsidered other research we had done during the same period, for example, research on how

“fun” versus “important” task instructions impact performance (Bianco et al., 2003), I began to realize that regulatory fit was a very general principle that applied to other orientations and strategies. Nonetheless, it was the specific work testing regulatory focus theory that led to the discovery of regulatory fit. Regulatory focus theory was the parent of regulatory fit theory, and as this new child developed it began to generate its own separate studies and findings (see Higgins, 2008a, 2009). At the end of this chapter, I will [p. 498 ↓] discuss how parent and child have benefited from one another.

Applicability to Social Issues

As I noted earlier, it was self-discrepancy theory that gave birth to regulatory focus theory. And my inspiration for self-discrepancy theory was wanting to understand the distinct psychological underpinnings of depression and anxiety disorders. My first collaborator on self-discrepancy theory, Tim Strauman, had the same inspiration as me. Indeed, while a graduate student at NYU, he earned two PhDs – one in social psychology and one in clinical psychology. After leaving NYU, he continued to wear both hats, including being a director of clinical training and helping clinically depressed and anxious clients as a therapist. He began to develop a new form of clinical psychotherapy based on self-discrepancy theory. After the birth of regulatory focus theory, he expanded and modified this therapy to take advantage of the new insights provided by regulatory focus theory.

Conceptual advances and empirical discoveries from developing regulatory focus theory since 1990 have increased psychologists' understanding of the differences between depression-related promotion failure and anxiety disorder-related prevention failure. This has led to Strauman and his collaborators developing and testing, in clinical trials, a new-generation psychotherapy called “self-system therapy” (Vieth et al., 2003). Interventions specifically designed to reduce the actual-ideal discrepancies of depressed patients have proven to be effective. Indeed, for a theory-specified subset of depressed patients, it has been shown to be even more effective than cognitive therapy (Strauman et al., 2006).

The emotional and motivational significance of the difference between a failure in the promotion system and a failure in the prevention system sheds new light regarding

other clinical phenomenon. There is evidence, for example, that among women who become a mother for the first time, having an actual-ideal discrepancy prior to the birth of their child predicts increased vulnerability to post-partum depression, whereas having an actual-ought discrepancy predicts *decreased* vulnerability to post-partum anxiety (Alexander and Higgins, 1993). There is also evidence that possessing an actual-ideal discrepancy is a vulnerability factor for *bulimic* eating disorders whereas possessing an actual-ought discrepancy is a vulnerability factor for *anorexic* eating disorders (Higgins et al., 1992; Strauman et al., 1991).

In addition to its implications for clinical phenomenon, regulatory focus theory has implications for interpersonal relations and intergroup relations as well. There is evidence that people are more willing to forgive another person who apologizes for hurting them and are more empathic concerning another person's suffering when the promotion or prevention nature of the forgiveness message or the other person's suffering fits the promotion or prevention focus of the recipient or perceiver (Houston, 1990; Santelli et al., 2009). This illustrates an interpersonal benefit from *similarity* in regulatory focus. There is evidence as well that *complementarity* in regulatory focus can also have interpersonal benefits. Recent research has found that long-term married partners with complementary regulatory focus orientations have higher relationship wellbeing (Bohns et al., 2009). What appears to be critical for such complementarity effects is for each partner to be able to assume a separate role on shared tasks (i.e., division of labor) so that each can use the goal strategy that fits their regulatory focus orientation, such as the promotion partner taking on the eager parts of the task and the prevention partner taking on the vigilant parts of the task. There is also evidence that responses to social exclusion vary in a manner that relates to regulatory focus, with individuals responding in a prevention manner when they are rejected but responding in a promotion [p. 499 ↓] manner when they are ignored (Molden et al., 2009).

Intergroup relations are also influenced by regulatory focus. Specifically, there is evidence that the classic phenomenon of *in-group favoritism* (see Levine and Moreland, 1998) varies by regulatory focus. Two separate research programs have shown that favoritism that rewards and embraces ingroup members is driven largely by promotion concerns, whereas favoritism that punishes and rejects outgroup members is driven largely by prevention concerns (Sassenberg et al., 2003; Shah et al., 2004) – *promoting us* versus *preventing them*. And this effect is evident even for subtle measures of

motivation for intergroup contact. For example, in a study by Shah et al. (2004), participants chose where to sit in a waiting room that had a backpack on a chair which supposedly was owned either by their future partner in an upcoming task or their future opponent. Participants with a stronger promotion focus chose to sit *closer to their teammate*, whereas a stronger prevention focus had no relation to sitting closer to one's teammate. In contrast, participants with a stronger prevention focus chose to sit *further away from their opponent*, whereas a stronger promotion focus had no relation to sitting further away from one's opponent.

Regulatory focus has other implications for intergroup relations as well. Being discriminated against is painful. But the nature of this pain and reactions to it can depend on perceivers' regulatory focus. When discrimination is perceived as blocking opportunities for advancement, the pain would involve dejection which has low motivational intensity. In contrast, when discrimination is perceived as a threat to one's security, the pain would involve agitation which has high motivational intensity. There is evidence, for example, that a prevention focus leads to more anger and agitation after social discrimination than a promotion focus, and especially when social discrimination is based on losses rather than on nongains (Sassenberg and Hansen, 2007). These differences in emotions and motivation could translate into people responding differently when they are discriminated against. Consistent with this, Quinn and Olson (2004) demonstrated that, compared to promotion-focused women, prevention-focused women report stronger intentions to engage in future behaviors that are aimed at reducing discrimination toward women, such as participating in protests on women discrimination issues, as well as reporting that they have performed such actions more frequently in the past. Interestingly, when behaviors that protest discrimination are explicitly framed in terms of removing obstacles to advancement, that is, removing a barrier to accomplishing progress, then promotion-focused women report stronger intentions to engage in such behaviors than do prevention-focused women.

There is also evidence that regulatory focus is relevant for reducing the negative impact of stereotype threat on performance (Steele et al., 2002). Keller (2007) has shown that if a promotion focus, rather than a prevention focus, can be induced under stereotype threat conditions, the negative impact of stereotype threat can be reduced. Keller (2007) argues that when individuals are in a promotion focus, stereotype threat is more likely to be experienced as a challenge rather than as a threat, which in turn creates

greater eagerness and engagement in maximal goals that enhance performance. Research by Seibt and Förster (2004) suggests that negative stereotypes induce a prevention focus that, in turn, motivates people to use vigilant strategies on a task (cf. Förster et al., 2004). If the task is one in which a vigilant strategy is useful, such as an analytical task, then this will not be a problem. But if the task requires the use of an eager strategy, or a mix of vigilant and eager strategies, then the prevention focus induced by negative stereotypes will hurt performance. On such tasks, Keller's (2007) intervention of inducing a promotion focus under stereotype threat conditions could be especially important.

[p. 500 ↓] One final way in which regulatory focus theory has been applied to social issues should be mentioned. Over the last several years, regulatory focus theory – often combined with the principle of regulatory fit – has been used to increase the effectiveness of persuasive messages (for reviews, see Cesario et al., 2008; Lee and Higgins, 2009). This application can be used, and *has* been used, to enhance the effectiveness of health messages. For instance, several studies have demonstrated that when recipients who are either promotion-focused or prevention-focused are given messages that are framed, respectively, in promotion-eager terms or in prevention-vigilant terms, the recipients are more persuaded to increase their consumption of fruits and vegetables (Cesario et al., 2004; Latimer et al., 2007; Spiegel et al., 2004), to use sunscreen (Keller, 2006; Lee and Aaker, 2004), increase physical activity (Latimer et al., 2008), and reduce intentions to smoke (Kim, 2006; Zhao and Pechmann, 2007).

In an early demonstration of this persuasion technique, Spiegel et al. (2004) gave participants health messages that advocated pursuit of the same desired end-state – eating more fruits and vegetables. The key manipulations took place as part of the messages that participants received. Although all participants received the same message advocacy (“eat more fruits and vegetables”), a promotion versus prevention focus was manipulated through the concerns that were highlighted within the messages – accomplishments for promotion and safety for prevention. Within each regulatory focus condition, participants were asked either to imagine the benefits they would get if they complied with the health message (eager strategy) or the costs they would incur if they didn't comply with the health message (vigilant strategy). The participants in the fit conditions (promotion recipients/eager message; prevention recipients/vigilant message) ate more fruits and vegetables in the week following the first session than

participants in the nonfit conditions (prevention recipients/eager message; promotion recipients/vigilant message). Latimer et al. (2008) extended these findings to show that a single message framed to fit individuals' chronic regulatory focus led to greater fruit and vegetable consumption even four months after message delivery.

Conclusion

From self-discrepancy theory to regulatory focus theory to regulatory fit theory, this *family of theories* has been developed and applied for over 20 years now. I should note that when a parent has a child, the parent does not stop developing. The child develops, but so too does the parent. Regulatory focus theory, for example, continued to develop on its own after giving birth to regulatory fit theory, as illustrated by the recent regulatory focus theory distinction between eager and vigilant *strategies* versus risky and conservative *tactics* (e.g., Scholer et al., 2010; see Scholer and Higgins, *in press*).

Importantly, it is not only the parent that affects the development of the child. The child affects the development of the parent as well. For example, regulatory fit theory taught regulatory focus theory that a particular strategy can have a consistent and stable association with a specific regulatory concern, such as eager with promotion and vigilant with prevention, because the strategy is *in the service of* sustaining that regulatory concern; a perspective that has proven useful when reconsidering the relation between culture and personality (Higgins, 2008b; Higgins, et al., *in press*). And in addition to a child affecting a parent, a grandchild can affect a grandparent. Regulatory fit theory, for example, provided new insights for self-discrepancy theory's understanding of anhedonia, that is, the inability to gain pleasure from normally pleasurable activities, which is a central symptom of the depression that is associated with severe actual-ideal discrepancies. Because an actual-ideal discrepancy [p. 501 ↓] is a promotion failure that reduces eagerness, and low eagerness is a nonfit for promotion, engagement in positive activities is weakened when people have severe actual-ideal discrepancies, which in turn deintensifies their attractiveness (see Higgins, 2006b). In this way, positive activities in general lose their attractiveness.

In my family of theories, grandparent, parent, child, and grandchild have all enriched one another. It has been an exciting journey, and I look forward to observing and

participating in further developments. I am certain there will be new discoveries and new surprises. Let me conclude as I have in the past:

Children teach parents to appreciate life in new ways. Children help parents to discover new things about the world. Theories can too. Always remember to love your theory, enjoy your theory, and help it develop. It is what makes life as a scientist a joy

(Higgins, 2006a).

References

Alexander, M.J. and Higgins, E.T. (1993) Emotional trade-offs of becoming a parent: How social roles influence self-discrepancy effects . Journal of Personality and Social Psychology , 65, 1259–1269. <http://dx.doi.org/10.1037/0022-3514.65.6.1259>

Atkinson, J.W. (1964) An Introduction to Motivation . Princeton: D. Van Nostrand.

Bargh, J.A., Lombardi, W.J. and Higgins, E.T. (1988) Automaticity of chronically accessible constructs in person x situation effects on person perception: It's just a matter of time . Journal of Personality and Social Psychology , 55, 599–605. <http://dx.doi.org/10.1037/0022-3514.55.4.599>

Bassili, J.N. (1996) Meta-judgmental versus operative indexes of psychological attributes: The case of measures of attitude strength . Journal of Personality and Social Psychology , 71, 637–653. <http://dx.doi.org/10.1037/0022-3514.71.4.637>

Bentham, J. (1988) The Principles of Morals and Legislation . Amherst, NY: Prometheus Books. (Original work published 1781).

Bianco, A.T., Higgins, E.T. and Klem, A. (2003) How ‘fun/importance’ fit impacts performance: Relating implicit theories to instructions . Personality and Social Psychology Bulletin , 29, 1091–1103. <http://dx.doi.org/10.1177/0146167203253481>

Bohns, V.K., Lucas, G., Molden, D.C., Finkel, E.J., Coolsen, M.K., Kumashiro, M., Rusbult, C.E. and Higgins, E.T. (2009) When opposites fit: Increased relationship strength from partner complementarity in regulatory focus . Unpublished manuscript, Columbia University.

Carver, C.S. and Scheier, M.F. (1981) Attention and Self-regulation: A Control-theory Approach to Human Behavior . New York: Springer-Verlag. <http://dx.doi.org/10.1007/978-1-4612-5887-2>

Carver, C.S. and Scheier, M.F. (1990a) Principles of self-regulation: Action and emotion . In E.T. Higgins, ed. and R.M. Sorrentino (eds), *Handbook of Motivation and Cognition: Foundations of Social Behavior* , 2, 3–52. New York: Guilford Press.

Carver, C.S. and Scheier, M.F. (1990b). Origins and functions of positive and negative affect: A control-process view . *Psychological Review* , 97, 19–35. <http://dx.doi.org/10.1037/0033-295X.97.1.19>

Cesario, J., Grant, H. and Higgins, E.T. (2004) Regulatory fit and persuasion: Transfer from ‘feeling right’ . *Journal of Personality and Social Psychology* , 86, 388–404. <http://dx.doi.org/10.1037/0022-3514.86.3.388>

Cesario, J., Higgins, E.T. and Scholer, A.A. (2008) Regulatory fit and persuasion: Basic principles and remaining questions . *Social and Personality Psychology Compass* , 2, 444–463. <http://dx.doi.org/10.1111/j.1751-9004.2007.00055.x>

Edwards, W. (1955) The prediction of decisions among bets . *Journal of Experimental Psychology* , 51, 201–214. <http://dx.doi.org/10.1037/h0041692>

Fazio, R.H. (1986) How do attitudes guide behavior? In R.M. Sorrentino, ed. and E.T. Higgins (eds), *Handbook of Motivation and Cognition: Foundations of Social Behavior* , pp. 204–243. New York: Guilford Press.

Festinger, L. (1957) *A Theory of Cognitive Dissonance* . Evanston: Row, Peterson.

Förster, J., Higgins, E.T. and Idson, C.L (1998) Approach and avoidance strength as a function of regulatory focus: Revisiting the ‘goal looms larger’ effect . *Journal of Personality and Social Psychology* , 75, 1115–1131.

Förster, J., Higgins, E.T. and Werth, L. (2004) How threat from stereotype disconfirmation triggers self-defense . *Social Cognition* , 22, 54–74.

Förster, J., Grant, H., Idson, L.C. and Higgins, E.T. (2001) Success/failure feedback, expectancies, and approach/avoidance motivation: How regulatory focus moderates classic relations . *Journal of Experimental Social Psychology* , 37, 253–260.

Freud, S. (1950) *Beyond the Pleasure Principle* . New York: Liveright. (Original work published 1920).

Freud, S. (1952) *A General Introduction to Psychoanalysis* . New York: Washington Square Press. (Original work published 1920).

Gollwitzer, P.M., ed. and Bargh, J.A. (eds) (1996) *The Psychology of Action: Linking Cognition and Motivation to Behavior* . New York: Guilford Press.

Gray, J.A. (1982) *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-hippocampal System* . New York: Oxford University Press.

Heider, F. (1958) *The Psychology of Interpersonal Relations* . New York: Wiley. <http://dx.doi.org/10.1037/10628-000>

Higgins, E.T. (1987) Self-discrepancy: A theory relating self and affect . *Psychological Review* , 94, 319–340. <http://dx.doi.org/10.1037/0033-295X.94.3.319>

Higgins, E.T. (1990) Personality, social psychology, and person-situation relations: Standards and knowledge activation as a common language . In L.A. Pervin (ed.), *Handbook of Personality* pp. 301–338. New York: Guilford Press.

Higgins, E.T. (1991) Development of self-regulatory and self-evaluative processes: Costs, benefits, and tradeoffs . In M.R. Gunnar, ed. and L.A. Sroufe (eds), *Self*

Processes and Development: The Minnesota Symposia on Child Psychology , 23, 125–165. Hillsdale, NJ: Erlbaum.

Higgins, E.T. (1996) Ideals, oughts, and regulatory focus: Affect and motivation from distinct pains and pleasures . In P.M. Gollwitzer, ed. and J.A. Bargh (eds), The Psychology of Action: Linking Cognition and Motivation to Behavior , pp. 91–114. New York: Guilford Press.

Higgins, E.T. (2000) Making a good decision: Value from fit . American Psychologist , 55, 1217–1230. <http://dx.doi.org/10.1037/0003-066X.55.11.1217>

Higgins, E.T. (2001) Promotion and prevention experiences: Relating emotions to nonemotional motivational states . In J.P. Forgas (ed.), Handbook of Affect and Social Cognition , pp. 186–211. Mahwah, NJ: Lawrence Erlbaum Associates.

Higgins, E.T. (2004) Making a theory useful: Lessons handed down . Personality and Social Psychology Review , 8, 138–145. http://dx.doi.org/10.1207/s15327957pspr0802_7

Higgins, E.T. (2006a) Theory development as a family affair . Journal of Experimental Social Psychology , 42, 549–552. <http://dx.doi.org/10.1016/j.jesp.2006.02.008>

Higgins, E.T. (2006b) Value from hedonic experience and engagement . Psychological Review , 113, 439–460. <http://dx.doi.org/10.1037/0033-295X.113.3.439>

Higgins, E.T. (2008a) Regulatory fit . In J.Y. Shah, ed. and W.L. Gardner (eds), Handbook of Motivation Science , pp. 356–372. New York: Guilford Press.

Higgins, E.T. (2008b) Culture and personality: Variability across universal motives as the missing link . Social and Personality Psychology Compass , 2. <http://dx.doi.org/10.1111/j.1751-9004.2007.00075.x>

Higgins, E.T. (2009) Regulatory fit in the goal-pursuit process . In G.B. Moskowitz, ed. and H. Grant (eds), The Psychology of Goals . New York: Guilford Press.

Higgins, E.T., Bond, R.N., Klein, R. and Strauman, T. (1986) Self-discrepancies and emotional vulnerability: How magnitude, accessibility, and type of discrepancy influence affect . *Journal of Personality and Social Psychology* 51, 5–15. <http://dx.doi.org/10.1037/0022-3514.51.1.5>

Higgins, E.T., King, G.A. and Mavin, G.H. (1982) Individual construct accessibility and subjective impressions and recall . *Journal of Personality and Social Psychology* , 43, 35–47. <http://dx.doi.org/10.1037/0022-3514.43.1.35>

Higgins, E.T., Klein, R. and Strauman, T. (1985) Self-concept discrepancy theory: A psychological model for distinguishing among different aspects of depression and anxiety . *Social Cognition* , 3, 51–76, <http://dx.doi.org/10.1521/soco.1985.3.1.51>

Higgins, E.T., Pierro, A. and Kruglanski, A.W. (in press) Re-thinking culture and personality: How self-regulatory universal create cross-cultural differences. In R.M. Sorrentino, ed. and S. Yamaguchi (eds), *Handbook of Motivation and Cognition Across Cultures* . New York: Academic Press.

Higgins, E.T., Rholes, W.S. and Jones, C.R. (1977) Category accessibility and impression formation . *Journal of Experimental Social Psychology* , 13, 141–154. <http://dx.doi.org/10.1016/S0022-1031%2877%2980007-3>

Higgins, E.T., Roney, C., Crowe, E. and Hymes, C. (1994) Ideal versus ought predilections for approach and avoidance: Distinct self-regulatory systems . *Journal of Personality and Social Psychology* , 66, 276–286. <http://dx.doi.org/10.1037/0022-3514.66.2.276>

Higgins, E.T., Shah, J. and Friedman, R. (1997) Emotional responses to goal attainment: Strength of regulatory focus as moderator . *Journal of Personality and Social Psychology* , 72, 515–525. <http://dx.doi.org/10.1037/0022-3514.72.3.515>

Higgins, E.T. and Silberman, I. (1998) Development of regulatory focus: Promotion and prevention as ways of living . In J. Heckhausen, ed. and C.S. Dweck (eds), *Motivation and Self-regulation Across the Life Span* , pp. 78–113. New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511527869.005>

Higgins, E.T. and Tykocinski, O. (1992) Self-discrepancies and biographical memory: Personality and cognition at the level of psychological situation . *Personality and Social Psychology Bulletin* , 18, 527–535. <http://dx.doi.org/10.1177/0146167292185002>

Higgins, E.T., Vookles, J. and Tykocinski, O. (1992) Self and health: How 'patterns' of self-beliefs predict types of emotional and physical problems . *Social Cognition* , 10, 125–150. <http://dx.doi.org/10.1521/soco.1992.10.1.125>

Houston, D.A. (1990) Empathy and the self: Cognitive and emotional influences on the evaluation of negative affect in others . *Journal of Personality and Social Psychology* , 59, 859–868. <http://dx.doi.org/10.1037/0022-3514.59.5.859>

Idson, L.C. and Higgins, E.T. (2000) How current feedback and chronic effectiveness influence motivation: Everything to gain versus everything to lose . *European Journal of Social Psychology* , 30, 583–592. <http://dx.doi.org/10.1002/1099-0992%28200007/08%2930:4%3C583::AID-EJSP9%3E3.0.CO;2-S>

Idson, L.C., Liberman, N. and Higgins, E.T. (2000) Distinguishing gains from nonlosses and losses from nongains: A regulatory focus perspective on hedonic intensity . *Journal of Experimental Social Psychology* , 36, 252–274. <http://dx.doi.org/10.1006/jesp.1999.1402>

Idson, L.C., Liberman, N. and Higgins, E.T. (2004) Imagining how you'd feel: The role of motivational experiences from regulatory fit . *Personality and Social Psychology Bulletin* , 30, 926–937. <http://dx.doi.org/10.1177/0146167204264334>

Kahneman, D., ed. , Diener, E., ed. and Schwarz, N. (eds) (1999) Well-being: The Foundations of Hedonic Psychology New York: Russell Sage Foundation.

Kahneman, D. and Tversky, A. (1979) Prospect theory: An analysis of decision under risk . *Econometrica* , 47, 263–291. <http://dx.doi.org/10.2307/1914185>

Keller, J. (2007) When negative stereotypic expectancies turn into challenge or threat: The moderating role of regulatory focus . *Swiss Journal of Psychology* , 66, 163–168. <http://dx.doi.org/10.1024/1421-0185.66.3.163>

Keller, J. (2008) On the development of regulatory focus: The role of parenting styles . European Journal of Social Psychology , 38, 354–364. <http://dx.doi.org/10.1002/ejsp.460>

Keller, P.A. (2006) Regulatory focus and efficacy of health messages . Journal of Consumer Research , 33, 109–114. <http://dx.doi.org/10.1086/504141>

Kim, Y. (2006) The role of regulatory focus in message framing for antismoking advertisements for adolescents . Journal of Advertising , 35, 143–151. <http://dx.doi.org/10.2753/JOA0091-3367350109>

Latimer, A.E., Rivers, S.E., Rench, T.A., Katulak, N.A., Hicks, A., Hodorowski, J.K., Higgins, E.T. and Salovey, P. (2008) A field experiment testing the utility of regulatory fit messages for promoting physical activity . Journal of Experimental Social Psychology , 44, 826–832. <http://dx.doi.org/10.1016/j.jesp.2007.07.013>

Latimer, A.E., Williams-Piehota, P., Katulak, N.A., Cox, A., Mowad, L.Z., Higgins, E.T. and Salovey, P. (in press) Promoting fruit and vegetable intake through messages tailored to individual differences in regulatory focus . Annals of Behavioral Medicine .

Lee, A.Y. and Aaker, J.L. (2004) Bringing the frame into focus: The influence of regulatory fit on processing fluency and persuasion . Journal of Personality and Social Psychology , 86, 205–218. <http://dx.doi.org/10.1037/0022-3514.86.2.205>

Lee, A.Y. and Higgins, E.T. (2009) The persuasive power of regulatory fit . In M. Wänke (ed.), Social Psychology of Consumer Behavior , pp. 319–333. New York: Psychology Press.

Levine, J.M. and Moreland, R.L. (1998) Small groups . In D.T. Gilbert, ed. , S.T. Fiske, ed. , and G. Lindzey (eds), The Handbook of Social Psychology , 4th Edition, pp. 415–469. Boston: McGraw-Hill.

Lewin, K. (1935) A Dynamic Theory of Personality New York: McGraw-Hill.

Lopes, L.L. (1987) Between hope and fear: The psychology of risk . In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* , 20, 255–295. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960416-5>

Manian, N., Papadakis, A.A., Strauman, T.J. and Essex, M.J. (2006) The development of children's ideal and ought Self-Guides: Parenting, temperament, and individual differences in guide strength . *Journal of Personality* , 74, 1619–1645. <http://dx.doi.org/10.1111/j.1467-6494.2006.00422.x>

Miller, N.E. (1959) Liberalization of basic S-R concepts: Extensions to conflict behavior, motivation, and social learning . In S. Koch (ed.) *Psychology: A Study of a Science*, Vol. 2: General Systematic Formulations, Learning, and Special Processes , pp. 196–292. New York: McGraw-Hill.

Molden, D.C., Lucas, G.M., Gardner, W.L., Dean, K. and Knowles, M.L. (2009) Motivations for prevention or promotion following social exclusion: Being rejected versus being ignored . *Journal of Personality and Social Psychology* 96, 415–431. <http://dx.doi.org/10.1037/a0012958>

Moretti, M.M. and Higgins, E.T. (1999a) Internal representations of others in self-regulation: A new look at a classic issue . *Social Cognition* , 17, 186–208. <http://dx.doi.org/10.1521/soco.1999.17.2.186>

Moretti, M.M. and Higgins, E.T. (1999b) Own versus other standpoints in self-regulation: Developmental antecedents and functional consequences . *Review of General Psychology* , 3, 188–223. <http://dx.doi.org/10.1037/1089-2680.3.3.188>

Mowrer, O.H. (1960) *Learning Theory and Behavior* . New York: John Wiley. <http://dx.doi.org/10.1037/10802-000>

Quinn, K.A. and Olson, J.M. (2004) Preventing prejudice or promoting progress? Regulatory framing and collective action . Unpublished manuscript.

Roney, C.J.R., Higgins, E.T. and Shah, J. (1995) Goals and framing: How outcome focus influences motivation and emotion . *Personality and Social Psychology Bulletin* , 21, 1151–1160. <http://dx.doi.org/10.1177/01461672952111003>

Santelli, A.G., Struthers, C.W. and Eaton, J. (2009) Fit to forgive: Exploring the interaction between regulatory focus, repentance, and forgiveness . *Journal of Personality and Social Psychology* , 96, 381–394. <http://dx.doi.org/10.1037/a0012882>

Sassenberg, K. and Hansen, N. (2007) The impact of regulatory focus on affective responses to social discrimination . *European Journal of Social Psychology* , 37, 421–444. <http://dx.doi.org/10.1002/ejsp.358>

Sassenberg, K., Kessler, T. and Mummendy, A. (2003) Less negative = more positive? Social discrimination as avoidance or approach . *Journal of Experimental Social Psychology* , 39, 48–58. <http://dx.doi.org/10.1016/S0022-1031%2802%2900519-X>

Scholer, A.A. and Higgins, E.T. (in press) Distinguishing levels of approach and avoidance: An illustration using regulatory focus theory. In A.J. Elliot (ed.), *Handbook of Approach and Avoidance Motivation* .

Scholer, A.A., Zou, X., Fujita, K., Stroessner, S.J. and Higgins, E.T. (2010) When risk-seeking becomes a motivational necessity . *Journal of Personality and Social Psychology* 99, 215–231. <http://dx.doi.org/10.1037/a0019715>

Seibt, B. and Förster, J. (2004) Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci . *Journal of Personality and Social Psychology* , 87, 38–56. <http://dx.doi.org/10.1037/0022-3514.87.1.38>

Shah, J. (2003) The motivational looking glass: How significant others implicitly affect goal appraisals . *Journal of Personality and Social Psychology* , 85, 424–439. <http://dx.doi.org/10.1037/0022-3514.85.3.424>

Shah, J. and Higgins, E.T. (2001) Regulatory concerns and appraisal efficiency: The general impact of promotion and prevention . *Journal of Personality and Social Psychology* , 80, 693–705. <http://dx.doi.org/10.1037/0022-3514.80.5.693>

Shah, J., Higgins, E.T. and Friedman, R. (1998) Performance incentives and means: How regulatory focus influences goal attainment . *Journal of Personality and Social Psychology* , 74, 285–293. <http://dx.doi.org/10.1037/0022-3514.74.2.285>

Shah, J.Y., Brazy, P.C. and Higgins, E.T. (2004) Promoting us or preventing them: Regulatory focus and manifestations of intergroup bias . *Personality and Social Psychology Bulletin* , 30, 433–446. <http://dx.doi.org/10.1177/0146167203261888>

Spiegel, S., Grant-Pillow, H. and Higgins, E.T. (2004) How regulatory fit enhances motivational strength during goal pursuit . *European Journal of Social Psychology* , 34, 39–54. <http://dx.doi.org/10.1002/ejsp.180>

Steele, C.M., Spencer, S. and Aronson, J. (2002) Contending with group image: The psychology of stereotype and social identity threat . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 34, 379–440. New York: Academic Press.

Strauman, T.J. (1989) Self-discrepancies in clinical depression and social phobia: Cognitive structures that underlie emotional disorders? *Journal of Abnormal Psychology* , 98, 14–22. <http://dx.doi.org/10.1037/0021-843X.98.1.14>

Strauman, T.J. (1992) Self-guides, autobiographical memory, and anxiety and dysphoria: Toward a cognitive model of vulnerability to emotional distress . *Journal of Abnormal Psychology* 101, 87–95. <http://dx.doi.org/10.1037/0021-843X.101.1.87>

Strauman, T.J. and Higgins, E.T. (1987) Automatic activation of self-discrepancies and emotional syndromes: When cognitive structures influence affect . *Journal of Personality and Social Psychology* , 53, 1004–1014. <http://dx.doi.org/10.1037/0022-3514.53.6.1004>

Strauman, T.J., Vieth, A.Z., Merrill, K.A., Kolden, G.G., Woods, T.E., Klein, M.H., Papadakis, A.A., Schneider, K.L. and Kwapil, L. (2006) Self-system therapy as an intervention for self-regulatory dysfunction in depression: A randomized comparison with cognitive therapy . *Journal of Consulting and Clinical Psychology* , 74, 367–376. <http://dx.doi.org/10.1037/0022-006X.74.2.367>

Strauman, T.J., Vookles, J., Berenstein, V., Chaiken, S. and Higgins, E.T. (1991) Self-discrepancies and vulnerability to body dissatisfaction and disordered eating . *Journal of Personality and Social Psychology* , 61, 946–956. <http://dx.doi.org/10.1037/0022-3514.61.6.946>

Thaler, R.H. (1980) Toward a positive theory of consumer choice .

Journal of Economic Behavior and Organization , 1, 39–60. <http://dx.doi.org/10.1016/0167-2681%2880%2990051-7>

Thorndike, E.L. (1935) The Psychology of Wants, Interests, and Attitudes . New York: Appleton-Century-Crofts.

Zhao, G. and Pechmann, C. (2007) The impact of regulatory focus on adolescents' response to antismoking advertising campaigns . Journal of Marketing Research , 44, 671–687. <http://dx.doi.org/10.1509/jmkr.44.4.671>

<http://dx.doi.org/10.4135/9781446249215.n24>

Handbook of Theories of Social Psychology: Volume 1

A Model of Behavioral Self-Regulation

Contributors: Charles S. Carver & Michael F. Scheier

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "A Model of Behavioral Self-Regulation"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n25>

Print pages: 505-526

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n25>

[p. 505 ↓]

Chapter 24: A Model of Behavioral Self-Regulation

Charles S. Carver, ed. and Michael F. Scheier, ed.

Abstract

This chapter describes the evolution (or perhaps accretion) of a set of ideas bearing on the self-regulation of action and emotion. The ideas were drawn from many sources, eventually yielding a viewpoint in which goal-directed action is seen as reflecting a hierarchical set of feedback control processes, and the creation and reduction of affect are seen as reflecting another set of feedback processes. Also embedded in the model is the idea that confidence and doubt influence whether the person continues to struggle against adversity or gives up the goal that the adversity is threatening. The portion of the model devoted to affect is of particular interest in that it generates two positions that differ substantially from those deriving from other theories. The first is that both approach and avoidance give rise to both positive and negative feelings; the second is that positive affect leads to coasting, reduction in effort regarding the goal under pursuit. The recent interest in dual-process models, which distinguish between top-down goal pursuit and reflexive responses to cues of the moment, has caused us to re-examine some of our previous assumptions, considering the possibility that behavior is triggered in two distinct ways.

Introduction

This chapter outlines several aspects of a perspective we have adopted over an extended period of time concerning the self-regulation of behavior and emotion. This perspective is more about the structure of behavior than it is about the content of behavior. It represents a viewpoint on the metaphorical bones and tendons that underlie very diverse sorts of action. We believe it is a viewpoint that is generally compatible with many other theories that are described in this book, standing beside them rather than in place of them.

There are two respects in which the ideas described here differ from those described elsewhere in this book. First, these ideas may be less a “theory” than a “meta-theory,” a very general way of conceptualizing interwoven functions. It is a declaration of belief about certain aspects of how complex systems are organized. Second, we actually developed on our own very little of the viewpoint we are about to describe. With a few exceptions, most of what we have done is to bring together ideas that had been developed by other people for their own reasons, and applied them to phenomena that are of interest to personality and social psychologists.

The viewpoint outlined here has long been identified with the term *self-regulation* (Carver and Scheier, 1981), a term that means different things to different people. We use it to imply purpose, with self-corrective [p. 506 ↓] adjustments taking place as needed, which originate from within the person. These elements converge in the view that human behavior is a continuing process of moving toward (and sometimes away from) goal values. Some people ascribe to *self-regulation* the additional quality of restraining or overriding impulses (e.g., Baumeister and Vohs, 2004). We generally do not. We address issues of overriding impulses later in the chapter. When we do, we use the more restrictive term *self-control*.

Intellectual History

The intellectual history of this viewpoint on behavior is vast. It extends through the development of ideas about mechanical governors and computing machines (e.g.,

Ashby, 1940; Rosenblueth et al., 1943; Wiener, 1948), and homeostatic mechanisms within the body (Cannon, 1932). Its roots include the literature of expectancy-value models of motivation (e.g., Bandura, 1986; Feather, 1982; Rotter, 1954), and general systems theory (Ford, 1987; von Bertalanffy, 1968) – the idea that mechanisms with similar structural and functional properties operate at many levels of abstraction. A full depiction of this history would overwhelm the page limits of this chapter and is well beyond our reach.

What we can do, however, and what we have done in the pages that follow, is to situate these ideas in the historical context in which we came to them and began to work with them. It is nearly impossible to do this without writing about our own professional histories, because pursuit of this viewpoint has occupied much of our professional attention over a period of over 30 years. That, then, is the form taken by this chapter.

Self-Awareness and Cybernetics

We were graduate students together at the University of Texas in the early 1970s. Although we were in personality psychology, we (along with several others in our program) were nearly as interested in social psychology. At about that time, Robert Wicklund, a young faculty member in the social psychology program, and Shelley Duval, a graduate student there, developed a theory that caught our attention. It was called objective self-awareness theory (Duval and Wicklund, 1972).

Self-Awareness and Conformity to Standards

This theory had several facets, but what was most interesting to us at the time was its analysis of some of the circumstances leading a person's behavior to conform to situationally salient standards. The theory held that when a person's attention is directed inward to the self, it gravitates to self-aspects that can be evaluated by comparison with some standard of correctness. The theory's authors held that such comparison would generally reveal a discrepancy between the actual self-state and the salient

standard. This would yield negative self-evaluation and negative affect. This, in turn, was presumed to prompt the person to seek ways to move out of the state of being either aware of the discrepancy or aware of the self.

There are two potential ways to escape this negative state. One is to avoid stimuli that induce self-awareness. No awareness of the discrepancy, no aversiveness. The other is to change the present state so that it conforms to the standard. Without a discrepancy, an awareness of the self is not aversive. It was the second of these responses – adjustment of behavior to match the standard – that stuck in our minds, and which led to some of our first explorations in the lab.

There we found that increased self-awareness could cause research participants to become both less punitive and more punitive to others, depending on circumstances. In one study, it led young men to behave more chivalrously toward young women, [p. 507 ↓] punishing them less intensely for errors in a task (Scheier et al., 1974). In another case, it led participants to conform more closely to the “hint” dropped by the experimenter that stronger punishment would produce faster learning, the stated goal of the task (Carver, 1974).

These were by no means the first studies of how behavior shifts when attention is self-directed (see Duval and Wicklund, 1972), but they were formative for us. They helped lead us into the laboratory in a more serious way, and they helped solidify our impression that self-awareness could be a powerful and systematic influence on behavior. Of importance for the future direction of our thinking, we had also found that self-awareness does not just dampen behavior. We had found that self-awareness could lead in either of opposing directions: to more aggression in some circumstances, to less aggression in other circumstances. It induced shifting of behavior toward whatever standard was salient.

Duval and Wicklund's (1972) theory lit our first real path into human research. It highlighted several phenomena to which we returned repeatedly. The one just described is that self-focus can cause behavior to conform more closely to salient standards. Another (also described above, in less detail) is the idea that people would sometimes act to avoid or escape from self-focus when confronting a discrepancy. Noted only briefly by Duval and Wicklund was the idea that some standards are

prescriptive and others are proscriptive, with the latter being less directive than the former, requiring “only that particular points on personally relevant dimensions be avoided” (1972: 14). We eventually looked at all these phenomena, but ultimately did so with a different conceptual framing than Duval and Wicklund had used.

Cybernetics and Systems

Duval and Wicklund's (1972) theory was developed toward the end of a period in which many social psychological theories were heavily influenced by more general drive-based models of motivation. Examples are cognitive dissonance theory (Festinger, 1957) and reactance theory (Brehm, 1966). Such theories relied on aversive drives to serve as the motivational engine of behavior. In classic drive theory, a need state develops over time, leading to an internal tension state – drive – which is aversive. The person thus is motivated to do something to reduce that aversiveness. Without the aversiveness, there is no need for action.

In social psychological applications of this idea, specific drives were ascribed to specific kinds of internal mental states and drive-specific classes of responses were posited. In dissonance theory, the conflict between two incompatible cognitions was said to create an aversive drive that could be reduced by reconciling or obscuring the conflict. In reactance theory, experiencing a loss of freedom was said to create an aversive drive that could be reduced by reasserting the freedom. In the same way, self-awareness theory held that awareness of a discrepancy from a salient standard produced an aversive drive state that could be reduced either by acting to reduce the discrepancy or by avoiding the self-aware state.

Drive theories are not, however, the only ways of thinking about motivation or the forces that guide behavior. Not long after encountering self-awareness theory, one of us was exposed for the first time to the concepts of cybernetics, a viewpoint that is quite different from drive theory. Cybernetics had already had a (relatively brief) heyday in the 1950s and 1960s. Probably the best-known example of this viewpoint was an engaging book by Miller et al. (1960). This book brought into the psychological lexicon the acronym TOTE, standing for test-operate-test-exit. These functions depict the action

of a feedback loop, though they do so in a way that emphasizes a sequence of discrete steps rather than simultaneous occurrence of all the processes.

The logic of the TOTE unit – test (compare present state with a standard), operate (make [p. 508 ↓] a change in the present state), test (compare again to ensure the change had the effect intended), and exit (move to the next thing that needs doing) – has an inescapable resemblance to core processes in self-awareness theory: that when attention is directed to the self, there is a tendency to compare one's present state with a salient standard of appropriate behavior, and to shift the present state so that it more closely approximates the standard. There was a real difference in connotation, though. To Duval and Wicklund (1972), this set of functions concerns *self-evaluation*. To Miller et al. (1960), this set of functions describes *the structure of goal-directed behavior*. It describes the way all human goals are attained.

We were rather taken by the idea that a small set of functions described the structure of behavior. We also resonated to the idea that a great deal of human behavior could be planned and conducted without the need for aversive internal states. It felt intuitive to us that a great deal of behavior was done because it was fun, or because the actions led to interesting outcomes. The idea that self-awareness processes paralleled the structure of a feedback loop, a construct that had been useful in a variety of places other than personality and social psychology, was also exciting.

Hierarchical Organization

The ideas put forward by Miller et al. (1960) opened up a new avenue for exploration in literatures that were largely unfamiliar to us. It turned out that a number of people had proposed ideas with a similar character, over a fairly extended period (e.g., MacKay, 1956, 1966; Powers, 1973; for further review see Miller et al., 1960). Of particular interest, and of particular impact on our thinking, was a provocative statement published as we were becoming interested in feedback concepts. It was a book by William Powers (1973), in which he forcefully put forward a model of how human behavior may reflect a hierarchy of feedback processes in simultaneous operation.

His goal was very ambitious. He tried to account for how the nervous system created physical movements by which intentions and even abstract human values become expressed behaviorally. The construct that took center stage in this account was the feedback loop. Powers tried to map several layers of feedback processes to aspects of the nervous system. Knowledge of the nervous system has progressed substantially since 1973, of course, and parts of the picture that Powers created are doubtlessly contradicted by later evidence. However, viability of the core idea that feedback processes underlie organized action need not depend entirely on details.

In Powers (1973) we found several strong themes. First, perhaps even more than Miller et al. (1960), Powers made a compelling case for the idea that the feedback construct was fully up to the challenge of accounting for the complexity of behavior. Not as one loop, of course, but as an interwoven network of loops, dealing with regulation of diverse properties simultaneously.

Second, he argued more specifically that the feedback processes underlying behavior form a hierarchy of varying levels of abstraction, which could be characterized by their properties. He started from the lowest level – regulation of muscle-fiber tensions – and worked his way upward. Each level deals with discrepancies at its own level; the reference value for a given level is the output of the level immediately above it. When a person does a relatively abstract behavior (e.g., expressing kindness to an elderly neighbor by shoveling the snow off her walks), all levels below that level of abstraction are at work simultaneously. This was (and remains) a very interesting idea with a great many implications.

Whether this model accurately portrays the control of action or only provides a comfortable illusion, we found it useful in two ways. First, it provides the sense that it is plausible to posit a way for the kinds of intentions that social and personality [p. 509 ↓] psychologists are interested in to find their way into action. Second, it suggests a reason to attend to the literature of motor control (e.g., Rosenbaum et al., 2001). Some may see that literature as relevant only to exercise science and industrial psychology. We disagree. We believe that this literature also has things to tell personality and social psychologists. Nature is a miser and a recycler. It is very likely that principles embodied in movement control have more than just a little in common with principles that are embodied in higher mental functions (Rosenbaum et al., 2001).

A third contribution made by Powers was his particular view of how higher levels of control might be construed. Although he devoted very little attention (15 pages) to the three highest levels of control he argued for – those most relevant to the subject matter of personality and social psychology – he chose labels and wrote descriptions of these levels and their relationships to one another that are intuitively on-target and evocative of important concepts in our field.

Programs are organizations of behavior with choice points. They are clearly sequential and orderly (though the order can be quite flexible). They seem to require attention. If they are sufficiently well learned to have acquired an automatic runoff character, they are not programs, but a lower level of control that Powers called *sequences*. Programs are the level of the Powers hierarchy that most closely resembles Miller et al.'s TOTE construct, because of the sequencing of steps and subroutines that programs imply.

Principles, the level above programs, are roughly akin to values (Schwartz and Bilsky, 1990; Schwartz and Rubel, 2005). They are a basis for making the choices that programs entail, and they suggest certain programs to enter and to avoid. What Powers called *system concepts* are the coalesced essences of entities that imply certain principles and not others. One might think of the overall sense of ideal self as one example, the sense of an ideal relationship as another.

We have used these upper levels of the Powers hierarchy for decades as a conceptual heuristic for thinking about the organization of behavior. We never found the time to study the ideas empirically. However, these ideas have a certain amount in common with those of Vallacher and Wegner's (1987) action identification theory. This theory posits that people can identify any action at varying levels of abstraction. In identifying the action at that level, presumably they are also regulating it at that level. Within this framework, two complementary tendencies play out over time and circumstances: As the person becomes more adept at a behavior, less notice must be taken of its lower-level elements, and the person drifts to a higher-level construal of it. If the person encounters difficulties with the action at the level at which it is presently identified, the person is pulled to a lower-level construal of the action, to sort out the bits at the lower level.

The notion of hierarchical organization has many implications. A full treatment of those implications is unfortunately beyond the scope of this chapter. However, interested readers may find broader discussions elsewhere (e.g., Carver and Scheier, 1998, 1999a, 1999b; Powers, 1973).

Confidence and Doubt, Effort and Disengagement

The preceding section described our shift toward control-theory principles as a meta-theoretical grounding for interpreting self-awareness effects, and some of the complexity that ultimately came along with that shift. In this section we return to self-awareness theory and to another consequence of self-focused attention. As noted earlier, self-awareness was held to cause the person either to change the present state to fit the standard or to avoid self-focus, generally by withdrawing from or avoiding the situation in which self-focus was being created. The feedback model addresses the reducing [p. 510 ↓] of the discrepancy. What about the other effect?

Several studies had, in fact, already found that when discrepancies were created, people acted to avoid self-awareness manipulations or to leave the situation (for reviews, see Carver and Scheier, 1981; Duval and Wicklund, 1972). On the other hand, evidence also suggested that these effects require more than just a discrepancy. In one study, for example, a situation was set up in which a discrepancy was portrayed as being either flexible or inflexible (Steenbarger and Aderman, 1979). Only when the discrepancy was seen as inflexible was self-focus experienced as aversive and only in that case did subjects avoid the self-focusing situation.

Our view was that whether self-focus led to increased efforts to reduce discrepancies or to disengagement of effort and avoidance of self-focusing stimuli depended on whether the person was confident or doubtful about eventual success at reducing the discrepancy. Confidence should lead to renewed (or greater) effort; sufficient doubt should lead to giving up and withdrawal.

We suggested that there is a sort of psychological watershed on the confidence dimension, and that the character of subsequent behavior flows either to renewed effort or to disengagement. Subsequent studies (reviewed in Carver and Scheier, 1981; see also Carver, 2003b) supported this view. Self-awareness enhanced the efforts of persons who had favorable expectations of being able to attain goals and cut short the efforts of persons who had unfavorable expectations. The opposing effects of self-focus as a function of differences in confidence suggest that there is indeed a breakpoint on the confidence dimension, analogous to a watershed on a mountain ridge.

Expectancy Theories in the Zeitgeist

We were not, of course, the only ones to have observed the importance of expectancies. Many theorists have emphasized that theme over an extended period (e.g., Atkinson, 1964; Bandura, 1986; Feather, 1982; Rotter, 1954; Tolman, 1938; Vroom, 1964). Adoption of the expectancy construct as a way of addressing discrepancy reduction versus withdrawal linked our view of self-awareness phenomena to this tradition of expectancy-value models of motivation. Expectancy models vary in other ways, but in at least one respect their core argument is similar: confidence of success keeps the person engaged in the effort to succeed, and greater efforts tend to foster success; expectation of failure leads to not trying, and not trying often leads to failure.

Two models proposed at about that time resonated particularly well with this thinking: those of Klinger (1975) and Wortman and Brehm (1975). Both held that two regions of the range of expectancies form a dichotomy, with the resultant behavior falling into two categories: effort versus disengagement or withdrawal. Klinger (1975) contrasted commitment with disengagement. Wortman and Brehm (1975) contrasted the reassertion of control tied to reactance with the giving-up response tied to helplessness. Brehm's thinking would later evolve into a view in which people exert as much effort as needed to successfully complete a task, up to the point where success no longer seems worth the effort or no longer seems possible, at which point effort stops (Brehm and Self, 1989). These statements all imply that giving up is not merely lower effort; it is a shift from one class of response to another.

Giving up and Moving on

The fact that a tendency to disengage plays a role in human self-regulation is hard to deny. However, it is also hard to capture all the roles this idea plays in one or two statements. For example, is disengagement good or bad? On the one hand, disengagement is often a maladaptive or dysfunctional response. A person who gives up whenever encountering [p. 511 ↓] difficulty will never accomplish anything. Without continued struggle, it is often impossible to overcome obstacles. Some goals in life should not be given up easily, even if the struggle is hard and painful.

On the other hand, disengagement (at some level, at least) is also a necessity, a natural and indispensable part of self-regulation. If people are ever to turn away from efforts at unattainable goals, if they are ever to back out of blind alleys, they must be able to disengage, to give up and start over somewhere else. The importance of disengagement is particularly obvious with regard to concrete, low-level goals: people must be able to remove themselves from literal blind alleys and wrong streets, give up plans that have become disrupted by unexpected events, and spend the night in the wrong city if they have missed the last plane home.

Disengagement is also important, however, with regard to higher-level goals (Wrosch et al., 2003a). It is important to disengage and move on with life after the loss of a close relationship (e.g., Cleiren, 1993; Weiss, 1988). People sometimes must give up goals that are deeply embedded in the self if those values create too much conflict and distress in their lives (Pyszczynski and Greenberg, 1992). There often are goals of childhood that must be given up as it becomes apparent that they will never be realized (Baltes et al., 1979; Heckhausen and Schulz, 1995). Giving up thus is a double-edged sword. One of the thorniest questions in life is how to decide when to hang on and when to let go (Pyszczynski and Greenberg, 1992).

Another issue here is that certain kinds of decisions have simultaneous overtones of both continued effort and disengagement. Consider the scaling back of aspirations. Sometimes progress toward a goal is going poorly, expectancies of success are dim, and you want to quit. Rather than quit altogether, though, you abandon the more

demanding goal for a less demanding one (e.g., a struggling student stops thinking in terms of an A and starts thinking in terms of a C).

This is a kind of limited disengagement. The first goal is being given up at the same time as the lesser one is being adopted. Limited disengagement has an important positive consequence: it keeps you engaged in the general domain in which you'd wanted to quit. By scaling back the goal (giving up in a small way), you keep trying to move ahead (thus *not* giving up, in a larger way).

A potential problem with the limited-disengagement strategy stems from the fact that goals are often interrelated. It may be fine in principle to lower your grade aspiration from an A to a C. But if a high grade in this course is a prerequisite to another goal – say, admission to medical school – the limited disengagement works only temporarily. The same issue will likely recur later on, with respect to the broader goal to which this one leads. In some cases, this bind is not easily resolved. If medical school is your ultimate educational goal and your grades are bad, some rearrangement of the ultimate goal is going to be necessary.

In the broader scheme, giving up on unattainable goals has multiple positive consequences. It conserves energy rather than waste it in futile pursuit of the unattainable (Nesse, 2000). It also eventually readies the person to take up alternative goals (Klinger, 1975). Finally, emotional pain from lost goals seems to reflect the combination of remaining committed to them and yet being unable to move forward (Carver and Scheier, 1998; Pyszczynski and Greenberg, 1992; Wrosch et al., 2003b). Disengaging fully from them thus removes a source of negative feelings.

Expectancies and Feedback Loops

The incorporation of the expectancy construct made fairly good sense with respect to the analysis of self-awareness processes. It provided a reason why self-focus would lead in some cases to discrepancy reduction efforts and in other cases to behavioral withdrawal. It is a little harder, though, to fit the expectancy construct into the model of [p. 512 ↓] feedback loops. In some ways, these assumptions about expectancies and effort versus disengagement seem very ad hoc.

If one steps back from the hierarchy as a whole, however, the idea becomes more plausible. Powers (1973) had argued that a higher-level control loop operates by resetting the reference values of loops at the next lower level. Some kinds of resetting can be thought of as adjustments in level of aspiration (which we characterized above as a limited disengagement). Other kinds of resetting may be more complex, involving changes in entire programs of action that are being considered for enactment. If this program of action is not creating desired results at the higher level, resetting of lower-level goal may require abandoning that strategy altogether and trying a different one (there often being many ways to skin a cat).

Discrepancy Enlargement

Thus far we have talked about only those feedback processes that reduce discrepancies. There also exist feedback processes that enlarge discrepancies. These feedback loops are unstable. Unless overridden, they enlarge discrepancies without end. Some people believe that this kind of feedback is always problematic and dysfunctional (Powers, 1973). Others believe that positive loops are an important part of complex systems (DeAngelis et al., 1986; Maruyama, 1963; McFarland, 1971), but that in living systems (and other cases in which positive feedback is adaptive), the effect of this loop is limited in some way or other. There may be a natural endpoint (e.g., sexual arousal increases to the point of orgasm, thus ending the increase), or the discrepancy enlarging function may be constrained by a discrepancy reducing function.

One might view some discrepancy enlarging loops as avoidance processes. Examples of potential reference values for discrepancy-enlarging loops in social-personality psychology would include feared or disliked possible selves (Markus and Nurius, 1986; Ogilvie, 1987) and negative reference groups. These are values to be avoided. If a positive standard can be viewed as a goal (Miller et al., 1960), these standards might be thought of as anti-goals (Carver and Scheier, 1998). If comparison of the present state with this standard suggests that the discrepancy is small, an effort to enlarge the discrepancy may follow.

As Duval and Wicklund (1972) noted, such standards are generally not as directive as positive standards. If a prohibitive standard is at one end of a dimension of variability,

however, direction is thereby provided. As an example, Carver and Humphries (1981) recruited Cuban American students, and ascribed opinions to a group that is a negative reference group to them: the Castro government of Cuba. When asked to report their own opinions on the same issues, they took pains to differ from the opinions attributed to Cuban officials. Further, the tendency to do so was increased by self-focus. Thus, self-focus can increase discrepancy enlargement as well as discrepancy reduction (for more examples see Carver, 2003b; Carver and Scheier, 1998).

We have also suggested that social and personality psychology has examples of discrepancy-enlarging loops being constrained by discrepancy-reducing loops. This pattern seems represented in Higgins's (1996) concept of the ought self (Carver et al., 1999) and in Ryan and Deci's (2000) concept of introjected values. In both constructs, the initial impetus to behavior is the desire to avoid social sanction; a good way to avoid social sanction is to locate a socially approved value that is different from (or opposite to) the disapproved value, and move toward it. By homing in on the positive value, one simultaneously escapes the feared or disliked value.

Approach and Avoidance

As just suggested, the dual concepts of discrepancy reducing and discrepancy enlarging [p. 513 ↓] loops map onto the general form of approach and avoidance processes. Incentives are approached by systems that close discrepancies between present conditions and the incentives. Threats are avoided by systems that enlarge discrepancies between present conditions and the threats. The logic of feedback processes thus provides a way to construe this dichotomy among motivations.

The idea that behavior reduces to approach and avoidance tendencies is not new (e.g., Miller, 1944; Miller and Dollard, 1941), but it has re-emerged in recent years (e.g., Davidson 1998; Elliot, 2008). The idea that two sorts of feedback functions map onto these classes of motivations has led us to be more attentive to differences between approach and avoidance behavior (e.g., Carver and White, 1994). That has particular relevance for the next topic.

Affect

As described earlier, the view we had adopted on self-awareness effects did not include any assumption about aversive drive states. Yet it was clear that people do sometimes experience negative affect when experiencing self-awareness. This was most likely when the discrepancy between state and standard was relatively fixed – when there was doubt about being able to move forward. Further thought about these issues helped lead us to an elaboration of the model with which we were working. It led us to hazard a guess about the source of affect.

Origins

What is affect? Affect is positive or negative feelings. In many ways affect is the heart of emotion, though the term emotion often incorporates connotations of physiological changes that frequently accompany hedonic experiences. Affect pertains to one's desires and whether they are being met (Clore, 1994; Frijda, 1986; Ortony et al., 1988). But what is the internal mechanism by which feelings arise?

Many different kinds of answers to this question have been offered, ranging from neurobiological (e.g., Davidson 1992) to cognitive (Ortony et al., 1988). We proposed an answer that focused on what appear to be some of the functional properties of affect (Carver and Scheier, 1990, 1998, 1999a, 1999b). In suggesting this answer, we used feedback control once more as an organizing principle. Now, however, the control bears on a different quality.

We suggested that feelings arise as a consequence of a feedback loop that operates simultaneously with the behavior-guiding process and in parallel to it. We regard its operation as automatic. The easiest characterization of what this second process is doing is that it is checking on how well the first process (the behavior loop) is doing. The input for this second loop thus is the *rate of discrepancy reduction in the action system over time*. (We focus first on discrepancy-reducing loops, then consider enlarging loops.)

Consider a physical analogy. Action implies change between states. Difference between states is distance. The action loop thus controls the psychological analog of distance. If the affect loop assesses the action loop's progress, then the affect loop is dealing with the psychological analog of velocity, the first derivative of distance over time. To the degree that this analogy is meaningful, the input to the affect loop should be the first derivative over time of the input used by the action loop.

Input (how well you are doing) does not by itself create affect; a given rate of progress has different affective consequences in different contexts. We argued that this input is compared to a reference value (cf. Frijda, 1986, 1988), just as in other feedback loops. In this case, the value is an acceptable or expected rate of behavioral discrepancy reduction. As in other feedback loops, the comparison checks for deviation from [p. 514 ↓] the standard. If there is a discrepancy, an error is sensed and the output function changes.

We think the error signal in this loop is manifest in experience as affect, a sense of positive or negative valence. A rate of progress below the criterion yields negative affect. A rate high enough to exceed the criterion yields positive affect. If the rate is not distinguishable from the criterion, there is no valence. In essence, the argument is that feelings with positive valence mean you are doing better at something than you need to, and feelings with negative valence mean you are doing worse than you need to (for detail, including supporting evidence, see Carver and Scheier, 1998, Chapters 8 and 9). The absence of affect means being neither ahead nor behind.

We are not arguing for a deliberative thinking through of whether rate conforms to the criterion rate. We assume that the testing is continuous and automatic. Nor are we arguing for a deliberative thinking about what the affective valence means. We assume that the meaning (i.e., being ahead versus behind) is intrinsic to the affect's valence, which itself arises automatically.

One implication of this line of thought is that affects that might potentially exist regarding any given action should fall on a bipolar dimension. That is, it should be the case that affect can be positive, neutral, or negative for any given goal-directed action, depending on how well or poorly the action seems to be attaining the goal.

Reference Criterion

What determines the criterion? There doubtlessly are many influences. Further, the orientation that a person takes to an action can induce a different framing that may change the criterion (Brendl and Higgins, 1996). What is used as a criterion is probably quite flexible when the activity is unfamiliar. If the activity is very familiar, the criterion is likely to reflect the person's accumulated experience, in the form of an expected rate (the more experience you have, the more you know what is reasonable to expect). Whether "desired" or "expected" or "needed" is most accurate as a depiction of the criterion rate may depend greatly on the context.

The criterion can also change, sometimes readily, sometimes less so. The less experience the person has in a domain, the easier it is to substitute one criterion for another. We believe, however, that change in rate criterion in a relatively familiar domain occurs relatively slowly. Continuing overshoots result automatically in an upward drift of the criterion, continuing undershoots result in a downward drift (see Carver and Scheier, 2000). Thus, the system recalibrates over repeated events. A (somewhat ironic) consequence of such recalibration would be to keep the balance of a person's affective experiences (positive to negative, across a span of time) relatively similar, even if the rate criterion changes considerably.

Two Kinds of Action Loops, Two Dimensions of Affect

So far we have addressed only approach loops. The view just outlined was that positive feeling exists when a behavioral system is making more than adequate progress *doing what it is organized to do*. The systems addressed so far are organized to reduce discrepancies. Yet there seems no obvious reason why the principle should not apply to systems that enlarge discrepancies. If such a system is making rapid enough progress attaining its ends, there should be positive affect. If it is doing poorly, there should be negative affect.

That affects of both valences are possible seems applicable to both approach and avoidance. That is, both approach and avoidance have the potential to induce positive feelings (by doing well), and both have the potential to induce negative feelings (by doing poorly). But doing well at *approaching an incentive* is not quite the same experience as doing well [p. 515 ↓] at *moving away from a threat*. Thus there may be differences between the two positives, and between the two negatives.

Drawing on the work of Higgins (e.g., 1987, 1996), we argue for two bipolar dimensions of affect, one bearing on approach, the other on avoidance (Carver, 2001; Carver and Scheier, 1998). Approach-related affect includes such positive affects as elation, eagerness, and excitement, and also such negative affects as frustration, anger, and sadness (Carver, 2004; Carver and Harmon-Jones, 2009). Avoidance-related affect includes such positive affects as relief, serenity, and contentment (Carver, 2009) and such negative affects as fear, guilt, and anxiety.

Affect and Action: Two Facets of One Event in Time

This two-layered viewpoint implies a natural connection between affect and action. That is, if the input function of the affect loop is a sensed rate of progress in action, the output function of the affect loop must be a change in the rate of progress in that action. Thus, the affect loop has a direct influence on what occurs in the action loop.

Some changes in rate output are straightforward. If you are lagging behind, you try harder. Some changes are less straightforward. The rates of many “behaviors” are defined not by pace of physical action but in terms of choices among potential actions, or entire programs of action. For example, increasing your rate of progress on a project at work may mean choosing to spend a weekend working rather than playing with family and friends. Increasing your rate of being kind means choosing to do an act that reflects kindness when an opportunity arises. Thus, change in rate must often be translated into other terms, such as concentration, or allocation of time and effort.

The idea of two feedback systems functioning jointly is something we stumbled into. As it happens, however, this idea is quite common in control engineering (e.g., Clark, 1996). Engineers have long recognized that having two systems functioning together – one controlling position, one controlling velocity – permits the device they control to respond in a way that is both quick and stable, without overshoots and oscillations.

The combination of quickness and stability in responding is desirable in many of the devices engineers deal with. It is also desirable in people. A person with very reactive emotions is prone to overreact and oscillate behaviorally. A person who is emotionally unreactive is slow to respond even to urgent events. A person whose reactions are between those extremes responds quickly but without behavioral overreaction and oscillation.

For biological entities, being able to respond quickly yet accurately confers a clear adaptive advantage. We believe this combination of quick and stable responding is a consequence of having both behavior-managing and affect-managing control systems. Affect causes people's responses to be quicker (because this control system is time sensitive); as long as the affective system is not over-responsive, the responses are also stable.

Our focus here is on how affects influence behavior, emphasizing the extent to which they are interwoven. However, note that the behavioral responses that are linked to the affects also lead to *reduction of the affects*. We thus would suggest that the affect system is, in a very basic sense, self-regulating (cf. Campos et al., 2004). It is undeniable that people also engage in voluntary efforts to regulate their emotions (e.g., Gross, 2007; Ochsner and Gross, 2008), but the affect system does a good deal of that self-regulation on its own.

Affect Issues

There are at least two important ways in which this view of affect differs from other theories bearing on emotion. One difference [p. 516 ↓] concerns the dimensional structure of affect (Carver, 2001).

Bipolarity

In some theories (though not all) affects are seen as having underlying dimensionality (e.g., Watson et al., 1999). Our view holds that affect generated through approach has the potential to be either positive or negative and that affect generated through avoidance also has the potential to be either positive or negative. Most dimensional models, however, ascribe to an approach system affects with positive valence and ascribe to an avoidance system affects with negative valence (e.g., Cacioppo et al., 1999; Lang et al., 1990; Watson et al., 1999).

There is at least some support for our view. There is evidence, albeit limited, that positive feelings of calmness and relief (as situationally relevant) relate to avoidance motivation (Carver, 2009; Higgins et al., 1997). There is far more evidence linking sadness to approach failure (for reviews see Carver, 2004; Higgins, 1996). There is also a good deal of evidence linking the approach system to the negative affect of anger (Carver and Harmon-Jones, 2009). Although it is clear that diverse negative feeling qualities coalesce with one another in moods (Watson, 2009), the evidence does not make that case with regard to situation-specific affective responses.

This issue is important, because it has implications for any attempt to identify a conceptual mechanism underlying creation of affect. Theories positing two unipolar dimensions assume that greater activation of a system translates to more affect of that valence (or more potential for affect of that valence). If the approach system relates both to positive and to negative feelings, however, this direct transformation of system activation to affect is not tenable. A conceptual mechanism is needed that naturally addresses both valences within the approach function (and, separately, the avoidance function). The mechanism described here does so.

Counterintuitive Effect of Positive Affect

A second issue also differentiates this model from other views (Carver, 2003a). Recall our argument that affect reflects the error signal from a comparison in a feedback loop. If this is so, affect is a signal to adjust rate of progress. This would be true whether the

rate is above the mark or below it – that is, whether affect is positive or negative. For negative feelings, this is intuitive. The first response to negative feelings is usually to try harder. If the person tries harder – and if more effort (or better effort) increases progress – the negative affect diminishes or ceases.

For positive feelings, prediction is counterintuitive. In this model, positive feelings arise when things are going better than they need to. But the feelings still reflect a discrepancy (albeit a positive one), and the function of a negative feedback loop is to keep discrepancies small. Such a system is organized in such a way that it “wants” to see neither negative nor positive affect. Either quality (deviation from the standard in either direction) would represent an “error” and lead to a change in output that would eventually reduce it. This view argues that people who exceed the criterion rate of progress (and who thus have positive feelings) will automatically tend to reduce subsequent effort in this domain. They will “coast” a little – ease back.

Expending greater effort to catch up when behind, and coasting when ahead, are both presumed to be specific to the goal domain to which the affect is attached, usually the goal from which the affect arises in the first place. We are not arguing that positive affect creates a tendency to coast *in general*, but with respect to the activity producing the positive feelings. We should also be clear that we are talking about the current, ongoing episode of action. We are *not* arguing that positive affect makes people less likely to do the behavior later on.

Does positive affect lead to coasting? Evidence consistent with this idea has been [p. 517 ↓] reported by Mizruchi (1991), by Louro et al. (2007), and by Fulford et al. (2010). Thus far, however, the issue is relatively untested. Some are skeptical about this idea, because that it is hard to see why a process would be built in that limits positive feelings – indeed, dampens them. We see at least two bases for it. The first lies in a basic biological principle: it is adaptive not to spend energy needlessly. Coasting prevents this. Brehm has similarly argued that people engage only as much effort as is needed to accomplish a given task, and no more (Brehm and Self, 1989).

The second basis for such a process stems from the fact that people have multiple simultaneous concerns. Given multiple concerns, people do not optimize their outcome on any one of them, but “satisfice” (Simon, 1953) – do a good enough job on each to

deal with it satisfactorily. This permits them to handle the many concerns adequately, rather than just any one of them. Coasting facilitates satisficing. A tendency to coast virtually defines satisficing regarding that particular goal. A tendency to coast also fosters satisficing of a broader set of goals, by allowing easy shift to other domains at little or no cost (see Carver, 2003a, for detail).

Affects and Priority Management

This line of argument implicates positive feelings in a broad function that deserves further attention: the shifting from one goal to another as focal in behavior (Dreisbach and Goschke, 2004; Shallice, 1978). This basic and very important phenomenon is often overlooked. Many goals are typically under pursuit simultaneously, but only one has top priority at a given moment. People need to shield and maintain intentions that are being pursued (cf. Shah et al., 2002), but they also need to be able to shift flexibly among goals (Shin and Rosenbaum, 2002).

The issue of priority management was addressed very creatively many years ago by Simon (1967). He proposed that emotions are calls for reprioritization. He suggested that emotion arising with respect to a goal that is out of awareness eventually induces people to interrupt their behavior and give that goal a higher priority than it had. The stronger the emotion, the stronger is the claim that the unattended goal should have higher priority than the goal that is presently focal.

Simon's discussion focused on cases in which a nonfocal goal demands a higher priority and *intrudes* on awareness. By strong implication, his discussion dealt only with negative affect. However, there is another way for priority ordering to shift: the focal goal can *relinquish its place*. Perhaps positive feelings also pertain to reprioritization, but rather than a call for higher priority, they reflect *reduction* in priority. Positive affect regarding avoidance (relief or tranquility) indicates that a threat has dissipated, no longer requires so much attention, and can assume a lower priority. Positive feelings regarding approach (happiness, joy) indicate that an incentive is being attained and could temporarily be put on hold because you are doing so well; thus, this goal can assume a lower priority (see Carver, 2003a).

Priority Management and Feelings of Depression

One more aspect of priority management must be addressed, concerning the idea that some goals are best abandoned. As noted earlier, we have long held that sufficient doubt about goal attainment yields a tendency to disengage from effort, and even to disengage from the goal itself. This is certainly a kind of priority shift, in that the abandoned goal now has an even lower priority than it had before. How does this case fit the ideas described thus far?

This case seems to contradict Simon's (1967) view that negative affect is a call for higher priority. But there is an important difference between two classes of negative [p. 518 ↓] affect associated with approach (Carver, 2003a, 2004; for this discussion we disregard avoidance). Some of these affects coalesce around frustration and anger. Others coalesce around sadness, depression, and dejection. The former relate to an increase in priority, the latter to a decrease.

In describing our view on affect, we said that approach-related affects fall on a dimension. However, the dimension is not a simple straight line. Progress below the criterion creates negative affect, as the incentive slips away. Inadequate movement gives rise to frustration, irritation, and anger, prompting more effort to overcome obstacles and reverse the inadequate current progress. But efforts sometimes do not change the situation. Indeed, losses preclude movement forward. Now the feelings are sadness, depression, despondency, and hopelessness. Behaviors also differ in this case. The person tends to disengage from – give up on – further effort.

In the first case, feelings of frustration and anger are a call for an upgrade in priority, an increase in effort, a struggle to gain the incentive despite setbacks. In the second case, feelings of sadness and depression accompany *reduction* of effort and a downgrade in priority. As described earlier, both the upgrade and the downgrade have adaptive functions in the appropriate situations.

Changes in the Theoretical Landscape: Two Modes of Functioning

During the last two decades, changes have occurred in how people view cognition and action. The implicit assumption that behavior is generally managed in a top-down, directive way has been challenged. Questions have been raised about the role of consciousness in many kinds of action. Interest has arisen in the idea that the mind has both explicit and implicit representations. These various ideas have also influenced how we think about ideas we have been working with.

Two-Mode Models

Several literatures have developed around the possible existence of two modes of functioning (for reviews see Carver and Scheier, 2009a; Carver et al., 2008). In personality, Epstein (e.g., 1973, 1994) has long advocated such a view. He argues that people experience reality through two systems. What he calls the *rational* system operates mostly consciously, uses logical rules, is verbal and deliberative, and thus is fairly slow. The *experiential* system is intuitive and associative in nature. It provides a quick and dirty way of assessing and reacting to reality. It relies on salient information and uses shortcuts and heuristics. It functions automatically and quickly. It is considered to be emotional (or at least very responsive to emotions) and nonverbal.

The experiential system is presumably older and more primitive neurobiologically. It dominates when speed is needed (as when the situation is emotionally charged). The rational system evolved later, providing a more cautious, analytic, planful way of proceeding. Being able to operate in that way has important advantages, if there is sufficient time and freedom from pressure to think things through. Both systems are presumed to be always at work, jointly determining behavior, though the extent of each one's influence can vary by situation and disposition.

A model in many ways similar to this, but with different roots, was proposed by Metcalfe and Mischel (1999). Drawing on decades of work on delay of gratification, Metcalfe

and Mischel (1999) proposed that two systems influence restraint. One they called a “hot” system: emotional, impulsive, and reflexive. The other they called a “cool” system: strategic, flexible, slower, and unemotional. How people respond to difficult situations depends on which system is in charge.

[p. 519 ↓] There are also several two-mode theories in social psychology (Chaiken and Trope, 1999). The essence of such a view has existed for a long time in the literature of persuasion. Strack and Deutsch (2004) have recently extended this reasoning more broadly into the range of behavioral phenomena of interest to social psychologists. They proposed a model in which overt social behavior is a joint output of two simultaneous modes of functioning, which they termed *reflective* and *impulsive*. Again, differences in the systems' operating characteristics lead to differences in behavior. The reflective system anticipates the future, makes decisions on the basis of those anticipations, and forms intentions. It is planful and wide-ranging in its search for relevant information. It is restrained and deliberative. The impulsive system acts spontaneously when its schemas or production systems are sufficiently activated. It acts without consideration for the future or for broader implications or consequences of the action. This depiction is very similar in some ways to the ideas of Epstein (1973, 1994) and Metcalfe and Mischel (1999).

The idea that there are two modes of functioning can also be linked to the burgeoning literature on explicit and implicit motives, knowledge structures, and attitudes. There often is little or no relation between explicit (self-report) measures and implicit measures of the same construct. Although it is fairly easy to see why that might be the case for variables such as prejudice (given the social desirability issues involved), it is less obvious why it would be so for such variables as the self-concept. Two-mode models suggest a possibility (Beever, 2005; Fazio and Olson, 2003). Implicit measures assess only associative links between pairs of elements. Explicit measures are symbolic, products of deliberative processing. Implicit knowledge presumably accrues through associative learning; explicit knowledge presumably accrues through verbal, conceptual learning. Perhaps associative and conceptual sources of knowledge about the self (or the world) are more independent of one another than is often assumed. Thus, the two sources may not agree well with each other over time, leading to different results from implicit and explicit measures.

Consistent with this line of thought, a number of studies have found that both implicit and explicit measures do predict aspects of behavior, but typically different aspects. Explicit measures predict deliberated decisions and intentions; implicit measures predict relatively automatic actions, nonverbal behaviors, and primed word completions (Dovidio et al., 1997; Neumann et al., 2004).

Two-mode thinking has also been very influential in developmental psychology. Rothbart and her colleagues have argued for the existence of three temperament systems: two for approach and avoidance, and a third termed *effortful control* (e.g., Derryberry and Rothbart, 1997; Rothbart and Posner, 1985; Rothbart et al., 2000; see also Nigg, 2000). Effortful control concerns (in part) the ability to suppress approach when it is situationally inappropriate. Effortful control is superordinate to approach and avoidance temperaments. The label *effortful* conveys the sense that this is an executive, planful activity, entailing the use of cognitive resources beyond those needed to react impulsively. This view of effortful control has substantial resemblance to depictions of the deliberative mode of the two-mode models outlined in previous sections.

Hierarchicality Re-Examined

Thus, several sources of theory propose that the mind functions in two modes. Indeed, the sources described above are far from an exhaustive list. All of them promote the inference that a deliberative mode of functioning uses symbolic and sequential processing and thus is relatively slow. They also suggest that a more impulsive or reactive mode of functioning uses associationist processing and is relatively fast. Many of the theories suggest that the two modes are [p. 520 ↓] semiautonomous in their functioning, competing with each other to influence actions. Indeed, many also point to situational variables that influence which mode dominates at a given time.

These kinds of ideas have influenced how we construe the hierarchy of control that was proposed by Powers (1973). We said earlier that *programs* of action entail decisions. They seem to be managed top-down, using effortful processing. Planfulness, an element of programs, is also a common characterization of behavior managed by

the reflective system. It seems reasonable to map program-level control onto the deliberative, reflective mode of functioning.

In contrast to this deliberative quality, well-learned sequences occur in a relatively automatic stream once they are triggered. Sequences (and all lower levels of control) are inevitably called up during the execution of programs. However, it may be that sequences can also be triggered more autonomously. Sequences may respond to cues that trigger them simply by virtue of associations in memory. In such cases, the operating characteristics would seem akin to those of the reactive mode of functioning.

We have often noted that the level of control that is functionally superordinate can vary by situations and persons (e.g., Carver and Scheier, 1998, 1999a). That is, it is possible to imagine cases in which a person is behaving intentionally according to a principle (e.g., a moral or ethical value), and it is possible to imagine cases in which the person is behaving according to a plan or program. It is also possible, however, to imagine cases in which the person is acting impulsively and spontaneously, without regard to either principle or plan.

In making this case in the past, our emphasis typically was simply on how sequences and programs differed. Now we are inclined to wonder if this particular differentiation is not perhaps more important than we had previously realized. Perhaps we have underappreciated the extent to which lower levels of self-regulatory structures can be triggered autonomously and their outputs enter the stream of ongoing action without oversight from higher levels (Carver and Scheier, 2002), and potentially even in conflict with values at higher levels. This seems an important question for further exploration.

Self-Control: Impulse and Restraint

The idea that there are conflicts between longer-term and shorter-term goals is also part of a literature on self-control and self-control failure (e.g., Baumeister et al., 1994). This literature focuses on cases in which a person is both motivated to act and motivated to restrain that action. This is essentially the case that is the focus of work on children's effortful control, and it also resembles the logical structure of the delay of gratification

paradigm. A difference is that in the self-control literature the intent often is to delay indefinitely rather than temporarily.

Although the self-control situation is often portrayed as pitting longer- and shorter-term goals against each other, the preceding discussion suggests a somewhat different view. The self-control situation may be pitting the two modes of processing against each other. This refraining would be consistent with the literature on self-control failure, which tends to portray such failures as involving a relatively automatic tendency to act in one way, opposed by a planful effort to restrain that act. The action that is being inhibited is often characterized as an impulse, a desire that would automatically be translated into action unless it is controlled (perhaps because this action is habitual, perhaps because it is more primal). The restraint is presumed to be effortful, and to depend on limited resources. If the planful part of the mind is fully enough able to attend to the conflict, the person can resist the impulse. If not, the impulse is more likely to be expressed. This portrayal seems quite consonant with the two-mode models of functioning.

[p. 521 ↓]

Applicability to Social Issues

A broader purpose of this volume is to suggest how theories may bear on applied topics. Authors were asked to evaluate the applicability of the theory they described to understanding and solving social issues and problems. This is a particularly difficult task for us, partly because our minds seem not to work that way, and partly because of the very nature of the ideas we have been writing about. As personality psychologists, most of what comes to us when we think about these ideas is what is going on in the person's mind. If we construe the idea of "social issues" quite broadly, however, a couple of applications do come to mind.

The affect-related part of the model provides what we think are useful ideas for understanding the nature of human distress: the idea that distress follows from the perception of not reaching desired goals (or not avoiding threats) combined with a continuing commitment to those goals. Put more simply, the model points to the bind that is inherent in a commitment to the unattainable. Reducing distress sometimes

requires finding better ways to move forward, but it sometimes requires the person to abandon goals and values and adopt new ones. Unfortunately, with few exceptions the model does not provide clear guidance about which of these options will be more profitable in a given case. An exception is that pursuit of some goals pulls and tears at the fabric of the self, enlarging discrepancies at a higher level even while closing them at a lower level. Pursuit of those goals will inevitably be problematic. We believe these ideas are useful for conceptualizing (and potentially treating) debilitating distress.

Another place where the ideas under discussion have had applied implications (if not exactly implications for social problems) concerns the portion of the theory bearing on expectations for future outcomes. Although we did not say so earlier in the chapter, we have used that theoretical principle to conceptualize and measure the individual difference dimension of optimism versus pessimism. This variable has been studied in a good deal of research in health psychology and related areas (Carver and Scheier, 2009b). It turns out to have important implications for how people respond to adversity, both psychologically and physically (Rasmussen et al., 2009; Solberg Nes and Segerstrom, 2006). Confidence about one's future keeps one in the struggle to adapt and thrive, and leads to better outcomes. This application indicates how the ideas embedded in this model are relevant to the broad sweep of human wellbeing.

Acknowledgments

Preparation of this chapter was supported by grants from the National Cancer Institute (CA64710), the National Science Foundation (BCS0544617), and the National Heart, Lung, and Blood Institute (HL65111, HL65112, HL076852, and HL076858).

References

Ashby, W.R. (1940) Adaptiveness and equilibrium . *Journal of Mental Science* , 86, 478–483.

Atkinson, J.W. (1964) An Introduction to Motivation . Princeton, NJ: Van Nostrand.

Baltes, P.B., Cornelius, S.W. and Nesselroade, J.R. (1979) Cohort effects in developmental psychology . In J.R. Nesselroade, ed. and P.B. Baltes (eds), Longitudinal Research in the Study of Behavior and Development , pp. 61–87. New York: Academic Press.

Bandura, A. (1986) Social Foundations of Thought and Action: A Social Cognitive Theory . Englewood Cliffs, NJ: Prentice-Hall.

Baumeister, R.F., Heatherton, T.F. and Tice, D.M. (1994) Losing Control: Why People Fail at Self-regulation . San Diego: Academic Press.

Baumeister, R.F., ed. and Vohs, K.D. (eds) (2004) Handbook of Self-regulation: Research, Theory, and Applications . New York: Guilford Press.

Beevers, C.G. (2005) Cognitive vulnerability to depression: A dual-process model . Clinical Psychology Review , 25, 975–1002. <http://dx.doi.org/10.1016/j.cpr.2005.03.003>

Brehm, J.W. (1966) A Theory of Psychological Reactance . New York: Academic Press.

Brehm, J.W. and Self, E.A. (1989) The intensity of motivation . Annual Review of Psychology , 40, 109–131. <http://dx.doi.org/10.1146/annurev.ps.40.020189.000545>

Brendl, C.M. and Higgins, E.T. (1996) Principles of judging valence: What makes events positive or negative? Advances in Experimental Social Psychology , 28, 95–160. <http://dx.doi.org/10.1016/S0065-2601%2808%2960237-3>

Cacioppo, J.T., Gardner, W.L. and Berntson, G.G. (1999) The affect system has parallel and integrative processing components: Form follows function . Journal of Personality and Social Psychology , 76, 839–855. <http://dx.doi.org/10.1037/0022-3514.76.5.839>

Campos, J.J., Frankel, C.B. and Camras, L. (2004) On the nature of emotion regulation . Child Development , 75, 377–394. <http://dx.doi.org/10.1111/j.1467-8624.2004.00681.x>

Cannon, W.B. (1932) The Wisdom of the Body New York: Norton.

Carver, C.S. (1974) Facilitation of physical aggression through objective self-awareness . *Journal of Experimental Social Psychology* , 10, 365–370. <http://dx.doi.org/10.1016/0022-1031%2874%2990032-8>

Carver, C.S. (2001) Affect and the functional bases of behavior: On the dimensional structure of affective experience . *Personality and Social Psychology Review* , 5, 345–356. http://dx.doi.org/10.1207/S15327957PSPR0504_4

Carver, C.S. (2003a) Pleasure as a sign you can attend to something else: Placing positive feelings within a general model of affect . *Cognition and Emotion* , 17, 241–261. <http://dx.doi.org/10.1080/02699930302294>

Carver, C.S. (2003b) Self-awareness . In M.R. Leary, ed. and J. Tangney (eds), *Handbook of Self and Identity* pp. 179–196. New York: Guilford Press.

Carver, C.S. (2004) Negative affects deriving from the behavioral approach system . *Emotion* , 4, 3–22. <http://dx.doi.org/10.1037/1528-3542.4.1.3>

Carver, C.S. (2009) Threat sensitivity, incentive sensitivity, and the experience of relief . *Journal of Personality* 77, 125–138. <http://dx.doi.org/10.1111/j.1467-6494.2008.00540.x>

Carver, C.S. and Harmon-Jones, E. (2009) Anger is an approach-related affect: Evidence and implications . *Psychological Bulletin* , 135, 183–204. <http://dx.doi.org/10.1037/a0013965>

Carver, C.S. and Humphries, C. (1981) Havana daydreaming: A study of self-consciousness and the negative reference group among Cuban Americans . *Journal of Personality and Social Psychology* , 40, 545–552. <http://dx.doi.org/10.1037/0022-3514.40.3.545>

Carver, C.S., Johnson, S.L. and Joormann, J. (2008) Serotonergic function, two-mode models of self-regulation, and vulnerability to depression: What depression has in common with impulsive aggression . *Psychological Bulletin* , 134, 912–943. <http://dx.doi.org/10.1037/a0013740>

Carver, C.S., Lawrence, J.W. and Scheier, M.F. (1999) Self-discrepancies and affect: Incorporating the role of feared selves . *Personality and Social Psychology Bulletin* , 25, 783–792. <http://dx.doi.org/10.1177/0146167299025007002>

Carver, C.S. and Scheier, M.F. (1981) Attention and Self-regulation: A Control-theory Approach to Human Behavior . New York: Springer Verlag. <http://dx.doi.org/10.1007/978-1-4612-5887-2>

Carver, C.S. and Scheier, M.F. (1990) Origins and functions of positive and negative affect: A control-process view . *Psychological Review* , 97, 19–35. <http://dx.doi.org/10.1037/0033-295X.97.1.19>

Carver, C.S. and Scheier, M.F. (1998) On the Self-regulation of Behavior . New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139174794>

Carver, C.S. and Scheier, M.F. (1999a) Themes and issues in the self-regulation of behavior . In R.S. Wyer, Jr. (ed.), *Advances in Social Cognition* , 12, 1–105. Mahwah, NJ: Erlbaum.

Carver, C.S. and Scheier, M.F. (1999b) Several more themes, a lot more issues: Commentary en the commentaries . In R.S. Wyer, Jr. (ed.), *Advances in Social Cognition* , 12, 261–302. Mahwah, NJ: Erlbaum.

Carver, C.S. and Scheier, M.F. (2000) Scaling back goals and recalibration of the affect system are processes in normal adaptive self-regulation: Understanding ‘response shift’ phenomena . *Social Science and Medicine* , 50, 1715–1722. <http://dx.doi.org/10.1016/S0277-9536%2899%2900412-8>

Carver, C.S. and Scheier, M.F. (2002) Control processes and self-organization as complementary principles underlying behavior . *Personality and Social Psychology Review* , 6, 304–315. http://dx.doi.org/10.1207/S15327957PSPR0604_05

Carver, C.S. and Scheier, M.F. (2009a) Action, affect, multi-tasking, and layers of control . In J.P. Forgas, ed. , R.F. Baumeister, ed. , and D. Tice (eds), *The Psychology of Self-regulation* , pp. 109–126. New York: Psychology Press.

Carver, C.S. and Scheier, M.F. (2009b) Optimism . In M.R. Leary, ed. and R.H. Hoyle (eds), *Handbook of Individual Differences in Social Behavior* , pp. 330–342. New York: Guilford Press.

Carver, C.S. and White, T.L. (1994) Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales . *Journal of Personality and Social Psychology* , 67, 319–333. <http://dx.doi.org/10.1037/0022-3514.67.2.319>

Chaiken, S.L., ed. and Trope, Y. (eds) (1999) *Dual-process Theories in Social Psychology* . New York: Guilford Press.

Clark, R.N. (1996) *Control System Dynamics* . New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139163873>

Cleiren, M. (1993) *Bereavement and Adaptation: A Comparative Study of the Aftermath of Death* . Washington, DC: Hemisphere.

Clore, G.C. (1994) Why emotions are felt . In P. Ekman, ed. and R.J. Davidson (eds), *The Nature of Emotion: Fundamental Questions* , pp. 103–111. New York: Oxford University Press.

Davidson, R.J. (1992) Anterior cerebral asymmetry and the nature of emotion . *Brain and Cognition* , 20, 125–151. <http://dx.doi.org/10.1016/0278-2626%2892%2990065-T>

Davidson, R.J. (1998) Affective style and affective disorders: Perspectives from affective neuroscience . *Cognition and Emotion* , 12, 307–330. <http://dx.doi.org/10.1080/026999398379628>

DeAngelis, D.L., Post, W.M. and Travis, C.C. (1986) Positive Feedback in Natural Systems (Biomathematics, Vol. 15). Berlin: Springer-Verlag. <http://dx.doi.org/10.1007/978-3-642-82625-2>

Derryberry, D. and Rothbart, M.K. (1997) Reactive and effortful processes in the organization of temperament . *Development and Psychopathology* , 9, 633–652. <http://dx.doi.org/10.1017/S0954579497001375>

Dovidio, J.F., Kawakami, K., Johnson, C., Johnson, B. and Howard, A. (1997) On the nature of prejudice: Automatic and controlled processes . *Journal of Experimental Social Psychology* , 33, 510–540. <http://dx.doi.org/10.1006/jesp.1997.1331>

Dreisbach, G. and Goschke, T. (2004) How positive affect modulates cognitive control: Reduced perseveration at the cost of increased distractibility . *Journal of Experimental Psychology: Learning, Memory, and Cognition* , 30, 343–353. <http://dx.doi.org/10.1037/0278-7393.30.2.343>

Duval, S. and Wicklund, R.A. (1972) *A Theory of Objective Self-awareness* . New York: Academic Press.

Elliot, A. (ed.) (2008) *Handbook of Approach and Avoidance Motivation* . Mahwah, NJ: Erlbaum.

Epstein, S. (1973) The self-concept revisited: Or a theory of a theory . *American Psychologist* , 28, 404–416. <http://dx.doi.org/10.1037/h0034679>

Epstein, S. (1994) Integration of the cognitive and the psychodynamic unconscious . *American Psychologist* , 49, 709–724. <http://dx.doi.org/10.1037/0003-066X.49.8.709>

Fazio, R.H. and Olson, M.A. (2003) Implicit measures in social cognition research: Their meaning and use . *Annual Review of Psychology* , 54, 297–327. <http://dx.doi.org/10.1146/annurev.psych.54.101601.145225>

Feather, N.T. (1982) Expectations and Actions: Expectancy-value Models in Psychology . Hillsdale: Erlbaum.

Festinger, L. (1957) *A Theory of Cognitive Dissonance* . Evanston, IL: Row, Peterson.

Ford, D.H. (1987) *Humans as Self-constructing Living Systems: A Developmental Perspective on Behavior and Personality* . Hillsdale, NJ: Erlbaum.

Frijda, N.H. (1986) *The Emotions* . Cambridge: Cambridge University Press.

Frijda, N.H. (1988) The laws of emotion . American Psychologist , 43, 349–358. <http://dx.doi.org/10.1037/0003-066X.43.5.349>

Fulford, D., Johnson, S.L., Llabre, M.M. and Carver, C.S. (2010) Pushing and coasting in dynamic goal pursuit: Coasting is attenuated in bipolar disorder . Psychological Science , 21, 1021–1027. <http://dx.doi.org/10.1177/0956797610373372>

Gross, J.J. (ed.) (2007) Handbook of Emotion Regulation . New York: Guilford Press.

Heckhausen, J. and Schulz, R. (1995) A life-span theory of control . Psychological Review , 102, 284–304. <http://dx.doi.org/10.1037/0033-295X.102.2.284>

Higgins, E.T. (1987) Self-discrepancy: A theory relating self and affect . Psychological Review , 94, 319–340. <http://dx.doi.org/10.1037/0033-295X.94.3.319>

Higgins, E.T. (1996) Ideals, oughts, and regulatory focus: Affect and motivation from distinct pains and pleasures . In P.M. Gollwitzer, ed. and J.A. Bargh (eds), The Psychology of Action: Linking Cognition and Motivation to Behavior , pp. 91–114. New York: Guilford Press.

Higgins, E.T., Shah, J. and Friedman, R. (1997) Emotional responses to goal attainment: Strength of regulatory focus as moderator . Journal of Personality and Social Psychology , 72, 515–525. <http://dx.doi.org/10.1037/0022-3514.72.3.515>

Klinger, E. (1975) Consequences of commitment to and disengagement from incentives . Psychological Review , 82, 1–25. <http://dx.doi.org/10.1037/h0076171>

Lang, P.J., Bradley, M.M. and Cuthbert, B.N. (1990). Emotion, attention, and the startle reflex . Psychological Review , 97, 377–395. <http://dx.doi.org/10.1037/0033-295X.97.3.377>

Louro, M.J., Pieters, R. and Zeelenberg, M. (2007) Dynamics of multiple-goal pursuit . Journal of Personality and Social Psychology , 93, 174–193. <http://dx.doi.org/10.1037/0022-3514.93.2.174>

MacKay, D.M. (1956) Toward an information-flow model of human behavior . British Journal of Psychology 47, 30–43. <http://dx.doi.org/10.1111/j.2044-8295.1956.tb00559.x>

MacKay, D.M. (1966) Cerebral organization and the conscious control of action . In J.C. Eccles (ed.), Brain and Conscious Experience , pp. 422–445. Berlin: Springer-Verlag.

Markus, H. and Nurius, P. (1986) Possible selves . American Psychologist , 41, 954–969. <http://dx.doi.org/10.1037/0003-066X.41.9.954>

Maruyama, M. (1963) The second cybernetics: Deviation-amplifying mutual causal processes . American Scientist , 51, 164–179.

McFarland, D.J. (1971) Feedback Mechanisms in Animal Behavior . New York: Academic Press.

Metcalfe, J. and Mischel, W. (1999) A hot/cool-system analysis of delay of gratification: Dynamics of willpower . Psychological Review , 106, 3–19. <http://dx.doi.org/10.1037/0033-295X.106.1.3>

Miller, G.A., Galanter, E. and Pribram, K.H. (1960) Plans and the Structure of Behavior . New York: Holt, Rinehart, and Winston. <http://dx.doi.org/10.1037/10039-000>

Miller, G.E. and Wrosch, C. (2007) You've gotta know when to fold 'em: Goal disengagement and systemic inflammation in adolescence . Psychological Science , 18, 773–777. <http://dx.doi.org/10.1111/j.1467-9280.2007.01977.x>

Miller, N.E. (1944) Experimental studies of conflict . In J. McV. Hunt (ed.), Personality and the Behavior Disorders , 1, 431–465. New York: Ronald Press.

Miller, N.E. and Dollard, J. (1941) Social Learning and Imitation . New Haven: Yale University Press.

Mizruchi, M.S. (1991) Urgency, motivation, and group performance: The effect of prior success on current success among professional basketball teams . Social Psychology Quarterly , 54, 181–189. <http://dx.doi.org/10.2307/2786935>

Nesse, R.M. (2000) Is depression an adaptation? *Archives of General Psychiatry* , 57, 14–20. <http://dx.doi.org/10.1001/archpsyc.57.1.14>

Neumann, R., Hülsenbeck, K. and Seibt, B. (2004) Attitudes towards people with AIDS and avoidance behavior: Automatic and reflective bases of behavior . *Journal of Experimental Social Psychology* , 40, 543–550. <http://dx.doi.org/10.1016/j.jesp.2003.10.006>

Nigg, J.T. (2000) On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology as a working inhibition taxonomy . *Psychological Bulletin* , 126, 220–246. <http://dx.doi.org/10.1037/0033-2909.126.2.220>

Ochsner, K.N. and Gross, J.J. (2008) Cognitive emotion regulation: Insights from social cognitive and affective neuroscience . *Current Directions in Psychological Science* , 17, 153–158. <http://dx.doi.org/10.1111/j.1467-8721.2008.00566.x>

Ogilvie, D.M. (1987) The undesired self: A neglected variable in personality research . *Journal of Personality and Social Psychology* , 52, 379–385. <http://dx.doi.org/10.1037/0022-3514.52.2.379>

Ortony, A., Clore, G.L. and Collins, A. (1988) *The Cognitive Structure of Emotions* . New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511571299>

Powers, W.T. (1973) *Behavior: The Control of Perception* . Chicago: Aldine.

Pyszczynski, T. and Greenberg, J. (1992). Hanging On and Letting Go: Understanding the Onset, Progression, and Remission of Depression . New York: Springer-Verlag. <http://dx.doi.org/10.1007/978-1-4613-9191-3>

Rasmussen, H.N., Scheier, M.F. and Greenhouse, J.B. (2009) Optimism and physical health: A meta-analytic review . *Annals of Behavioral Medicine* , 37, 239–256. <http://dx.doi.org/10.1007/s12160-009-9111-x>

Rosenbaum, D.A., Carlson, R.A. and Gilmore, R.O. (2001) Acquisition of intellectual and perceptual-motor skills . *Annual Review of Psychology* , 52, 453–470. <http://dx.doi.org/10.1146/annurev.psych.52.1.453>

Rosenblueth, A., Wiener, N. and Bigelow, J. (1943) Behavior, purpose, and teleology . Philosophy of Science , 10, 18–24. <http://dx.doi.org/10.1086/286788>

Rothbart, M.K., Ahadi, S.A. and Evans, D.E. (2000). Temperament and personality: Origins and outcomes . Journal of Personality and Social Psychology , 78, 122–135. <http://dx.doi.org/10.1037/0022-3514.78.1.122>

Rothbart, M.K. and Posner, M. (1985) Temperament and the development of self-regulation . In L.C. Hartlage, ed. and C.F. Telzrow (eds), *The Neuropsychology of Individual Differences: A Developmental Perspective* , pp. 93–123. New York: Plenum.

Rotter, J.B. (1954) Social Learning and Clinical Psychology . New York: Prentice-Hall. <http://dx.doi.org/10.1037/10788-000>

Ryan, R.M. and Deci, E.L. (2000) Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being . American Psychologist , 55, 68–78. <http://dx.doi.org/10.1037/0003-066X.55.1.68>

Scheier, M.F., Fenigstein, A. and Buss, A.H. (1974) Self-awareness and physical aggression . Journal of Experimental Social Psychology , 10, 264–273. <http://dx.doi.org/10.1016/0022-1031%2874%2990072-9>

Schwartz, S.H. and Bilsky, W. (1990) Toward a theory of the universal content and structure of values: Extensions and cross-cultural replications . Journal of Personality and Social Psychology , 58, 878–891. <http://dx.doi.org/10.1037/0022-3514.58.5.878>

Schwartz, S.H. and Rubel, T. (2005) Sex differences in value priorities: Cross-cultural and multimethod studies . Journal of Personality and Social Psychology , 89, 1010–1028. <http://dx.doi.org/10.1037/0022-3514.89.6.1010>

Shah, J.Y., Friedman, R. and Kruglanski, A.W. (2002) Forgetting all else: On the antecedents and consequences of goal shielding . Journal of Personality and Social Psychology , 83, 1261–1280. <http://dx.doi.org/10.1037/0022-3514.83.6.1261>

Shallice, T. (1978) The dominant action system: An information-processing approach to consciousness . In K.S. Pope, ed. and J.L. Singer (eds), *The Stream of Consciousness*:

Scientific Investigations into the Flow of Human Experience , pp. 117–157. New York: Wiley.

Shin, J.C. and Rosenbaum, D.A. (2002) Reaching while calculating: Scheduling of cognitive and perceptual-motor processes . *Journal of Experimental Psychology: General* , 131, 206–219. <http://dx.doi.org/10.1037/0096-3445.131.2.206>

Simon, H.A. (1953) Models of Man . New York: Wiley.

Simon, H.A. (1967) Motivational and emotional controls of cognition . *Psychology Review* , 74, 29–39. <http://dx.doi.org/10.1037/h0024127>

Solberg Nes, L. and Segerstrom, S.C. (2006) Dispositional optimism and coping: A meta-analytic review . *Personality and Social Psychology Review* , 10, 235–251. http://dx.doi.org/10.1207/s15327957pspr1003_3

Steenbarger, B.N. and Aderman, D. (1979) Objective self-awareness as a nonaversive state: Effect of anticipating discrepancy reduction . *Journal of Personality* , 47, 330–339. <http://dx.doi.org/10.1111/j.1467-6494.1979.tb00206.x>

Strack, F. and Deutsch, R. (2004) Reflective and impulsive determinants of social behavior . *Personality and Social Psychology Review* , 8, 220–247. http://dx.doi.org/10.1207/s15327957pspr0803_1

Tolman, E.C. (1938) The determiners of behavior at a choice point . *Psychological Review* , 45, 1–41. <http://dx.doi.org/10.1037/h0062733>

Vallacher, R.R. and Wegner, D.M. (1987) What do people think they're doing? Action identification and human behavior . *Psychological Review* , 94, 3–15. <http://dx.doi.org/10.1037/0033-295X.94.1.3>

von Bertalanffy, L. (1968) General Systems Theory . New York: Braziller.

Vroom, V.H. (1964) Work and Motivation . New York: Wiley.

Watson, D. (2009) Locating anger in the hierarchical structure of affect: Comment on Carver and Harmon-Jones . Psychological Bulletin , 135, 205–208. <http://dx.doi.org/10.1037/a0014413>

Watson, D., Wiese, D., Vaidya, J. and Tellegen, A. (1999) The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence . Journal of Personality and Social Psychology , 76, 820–838. <http://dx.doi.org/10.1037/0022-3514.76.5.820>

Weiss, R.S. (1988) Loss and recovery . Journal of Social Issues , 44, 37–52. <http://dx.doi.org/10.1111/j.1540-4560.1988.tb02075.x>

Wiener, N. (1948) Cybernetics: Control and Communication in the Animal and the Machine . Cambridge, MA: MIT Press.

Wortman, C.B. and Brehm, J.W. (1975) Responses to uncontrollable outcomes: An integration of reactance theory and the learned helplessness model . In L. Berkowitz (ed.) Advances in Experimental Social Psychology , 8, 277–336. New York: Academic Press.

Wrosch, C., Scheier, M.F., Carver, C.S. and Schulz, R. (2003a) The importance of goal disengagement in adaptive self-regulation: When giving up is beneficial . Self and Identity , 2, 1–20 <http://dx.doi.org/10.1080/15298860309021>

Wrosch, C., Scheier, M.F., Miller, G.E., Schulz, R. and Carver, C.S. (2003b) Adaptive self-regulation of unattainable goals: Goal disengagement, goal re-engagement, and subjective well-being . Personality and Social Psychology Bulletin , 29, 1494–1508. <http://dx.doi.org/10.1177/0146167203256921>

<http://dx.doi.org/10.4135/9781446249215.n25>

Handbook of Theories of Social Psychology: Volume 1

Mindset Theory of Action Phases

Contributors: Peter M. Gollwitzer

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Mindset Theory of Action Phases"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n26>

Print pages: 526-547

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n26>

[p. 526 ↓]

Chapter 25: Mindset Theory of Action Phases

Peter M. Gollwitzer, ed.

Abstract

Mindset theory of action phases is based on the distinction between motivation and volition as proposed by the Rubicon model which claims that prior to crossing the Rubicon (i.e., making a goal decision) motivational principles apply whereas thereafter volitional principles set in. The latter are concerned with goal implementation, whereas the former relate to the choosing of goals. Mindset theory of action phases proposes that different cognitive procedures are activated when people tackle the task of choosing goals versus implementing them. The respective task demands determine the features that characterize the deliberative versus implemental mindset. These pertain to what type of information is preferably processed and how it is analyzed. Mindset research has produced findings that not only support the motivation versus volition distinction but also enlighten various debates and theories in social psychology (e.g., optimism versus realism debate, dual process theories, goal theory). Mindset theory of action phases has also spurred research on effective planning by pointing to implementation intentions (i.e., if-then plans). This research has had much applied impact. When it was linked up to research on strategies of motivationally smart goal setting (i.e., mental contrasting), it initiated the development of a time- and cost-effective behavior change intervention.

Introduction

During my graduate education in the late 1970s at the University of Texas at Austin, my mentor, Robert Wicklund, and I started to conceive of people's selves or identities as goals. We thought that people can very well set themselves goals to become a good parent, a brilliant scientist, or a great athlete. If one takes this perspective, the self of a person is no longer just something to understand (self-concept) and like (self-esteem) but something to be achieved (identity goal). We turned to the writings of Kurt Lewin (1926) and his students, whose tension system theory of goal pursuit, with its notion of substitution, was very helpful to developing our theory of symbolic self-completion (Gollwitzer and Kirchhof, 1998; Wicklund and Gollwitzer, 1982). The main proposition of self-completion theory is that once people have set themselves certain identity or self-defining goals, they respond to failure experiences, shortcomings, or barriers not with retreat but instead with intensified efforts to reach the goal. These efforts, however, do not [p. 527 ↓] have to alleviate the problem at hand but may involve resorting to any substitute that indicates goal attainment (e.g., showing off relevant status symbols, engaging in identity-relevant activities, describing oneself as having the required personal attributes; for recent research on self-completion theory see Gollwitzer et al., 2009; Harmon-Jones et al., 2009; Ledgerwood et al., 2007).

In the early 1980s, Heinz Heckhausen invited me to join the newly founded Max Planck Institute for Psychological Research at Munich to start a research unit called Motivation and Action. We quickly realized that we had a directly opposed conceptual view of motivation. Whereas Heinz Heckhausen's motivation was that of an expectancy-value theorist in the tradition of Atkinson (1957) and Heckhausen (1977) and was thus fueled by the perceived feasibility and desirability of a given action, my motivation was that of Lewin's (1926) tension system and was resting in the determination or commitment a person holds with respect to the action goal at hand. Apparently, in the research on self-completion (Wicklund and Gollwitzer, 1982) I had been studying issues of goal striving (i.e., thoughts and behavior directed toward existing goals), whereas Heinz Heckhausen in his work on achievement motivation (Heckhausen, 1977) had focused on issues of goal setting (i.e., what goals people find attractive and feasible, and thus choose for themselves).

The Rubicon Model of Action Phases

To highlight this insight, we suggested making a distinction between motivation and volition. Following the conceptual terms used by Lewin (1926) and Narziß Ach (1935), we dubbed the goal-striving-related motivation with the term *volition*, and kept the term *motivation* for the goal-setting-related motivation. More importantly, in an attempt to integrate these two kinds of phenomena (i.e., motivation and volition) we developed the Rubicon model of action phases (Heckhausen, 1987; Heckhausen and Gollwitzer, 1987). This model suggests that the course of action can be segmented into four different, consecutive phases that differ in terms of the tasks that are to be solved by the individual given that s/he wants to execute a given course of action successfully. The first phase (predecision phase) is said to pose the task of setting preferences among wishes and desires by deliberating their desirability and feasibility. As people's motives and needs produce more wishes and desires than can possibly be realized, the individual is forced to choose among these desires and by doing so turn them into goals. Once goals are set (i.e., the Rubicon has been crossed), the individual faces the second task (preaction phase), which is getting started with goal-directed behaviors. This may be simple if the necessary goal-directed actions are well practiced and routine but complex if the individual is still undecided about where, when, and how to act. In such complex cases, the execution of goal-directed action has to be planned by deciding on when, where, and how to act. The third task (action phase) is bringing the initiated goal-directed action to a successful ending, and this is best achieved by determined and persistent pursuit of goal completion. Finally, in the fourth task (postaction phase), the individual needs to decide whether the desired goal has indeed been achieved or whether further striving is needed.

The Rubicon model of action phases postulates that a person's psychological functioning in each of these phases is governed by different principles. Classic theories of motivation (adhering to the restricted definition of motivation as determined by feasibility and desirability; Atkinson, 1957; Feather and Newton, 1982; Heckhausen, 1977) are said to be well suited to explicate the psychological processes associated with the predecision and postaction phases, whereas theories of volition (i.e., theories on the self-regulation of goal attainment; Lewin, 1926; [p. 528 ↓] Mischel, 1974;

Mischel and Patterson, 1978) are most appropriate to explaining the psychological processes that characterize the preaction and action phase. In other words, the predecision and postaction phases are expected to encompass motivational phenomena and processes in the classic sense of the term, whereas in the phases in between volitional phenomena and processes are thought to occur.

This radical statement needed empirical support, and therefore Heckhausen and I conducted an early experiment aimed at demonstrating that individuals placed in the predecision phase evidence different cognitive functioning than do individuals in the preaction phase (Heckhausen and Gollwitzer, 1987, Study 2). Assuming that deliberation of the desirability and feasibility of wishes and desires (the task of the predecision phase) is cognitively more demanding than committing to a plan that specifies, when, where, and how one wants to perform goal-directed actions (the task of the preaction phase), we expected that deliberating individuals experienced a higher cognitive load than planning (i.e., preaction) individuals. We therefore interrupted experimental participants who were either in the middle of deliberating a choice between two different tests that presumably measured their creative potential or in the middle of planning how to perform the test they had just chosen and then asked them to take a short-term memory test (i.e., a noun span test that presented nouns irrelevant to the creativity tests at hand). We expected that deliberating participants, because of heightened cognitive load, would evidence a reduced noun span, compared with their span as measured at the beginning of the experiment. We also expected that deliberating participants would evidence a comparatively more reduced noun span than planning participants because laying down a plan on how to act was expected to take up less cognitive resources than deliberating the pros and cons of a goal decision.

To our surprise, the results were just opposite to what we had predicted (Heckhausen and Gollwitzer, 1987, Study 2). The deliberating participants showed an increase in their short-term memory capacity, compared with both their own prior span and the span of the planning participants. In an effort to reduce our confusion about these unexpected findings, I turned to Gerhard Strube, at the time a cognitive psychologist at the Max-Planck-Institute for Psychological Research, and he pointed me to the classic concept of mindset as originally advanced at the turn of the century by the German psychologists Külpe (1904), Marbe (1915), Orth (1903), and Watt (1905), all members of the Würzburg school. These early cognitive psychologists had discovered

that becoming intensively involved with performing a given task activates exactly those cognitive procedures that help task completion. The created mindset (i.e., the sum total of the activated cognitive procedures) is the cognitive orientation most conducive to successful task performance.

The mindset notion allows interpreting the observed noun span data as follows: deliberating between potential action goals activates cognitive procedures (*the deliberative mindset*) that facilitate the task of the predecision phase, which is to set preferences. As undecided individuals do not know yet in which direction their decisions will finally take them, a heightened receptiveness to all kinds of information (open-mindedness) seems appropriate and functional to task solution. Similarly, planning out the implementation of a chosen goal should activate cognitive procedures (*the implemental mindset*) that facilitate the task of the preaction phase (i.e., getting started on the chosen goal). As this requires a more focused and selective orientation to processing information, closed-mindedness rather than open-mindedness with respect to available information seems called for. This postulated difference in receptiveness between deliberating and planning individuals is expressed in the fact that the experimental participants in the Heckhausen and Gollwitzer (1987, Study 2) noun span study processed the presented information in the noun span task [p. 529 ↓] faster than planning participants (i.e., the deliberating participants demonstrated a broader noun span than the planning participants).

Mindset Theory of Action Phases

However, isn't all of this post hoc? This is exactly the type of worry that made me use the mindset notion as a hypothesis-generating device for subsequent research and thus for developing a comprehensive mindset theory of action phases in my *Habilitationsschrift* (i.e., a second, more extensive doctoral thesis that in Germany is a prerequisite for attaining a tenured professorship; Gollwitzer, 1987). A summary of this thesis (Gollwitzer, 1990) can be found in a chapter in *Motivation and Cognition* edited by E. Tory Higgins and Richard M. Sorrentino (1990), and a more extensive version in a German book, *Abwägen und Planen* [Deliberating and Planning] (Gollwitzer, 1991). If one analyzes the unique demands of the task of choosing between wishes and desires in the predecision phase versus the typical demands of the task of getting

started on a chosen goal in the preaction phase, it becomes possible to detect further cognitive features of the deliberative as compared to the implemental mindset that can then be tested in new experiments. The task of deliberating in the predecisional phase is to choose, from among various wishes and desires, those few that one wants to realize (Gollwitzer, 1990). The criteria for selection should be the feasibility and desirability of the wishes and desires at issue. The systematic analysis of the chances of realization as well as the desirability of realization requires that relevant information be preferentially encoded and retrieved. But such cognitive tuning to this information should not suffice, as feasibility-related information needs to be analyzed objectively (and not in a self-serving manner), and desirability-related information needs to be analyzed in an impartial manner (and not in a biased manner). Only if feasibility-related information is analyzed realistically, and the pros and cons are weighed impartially, can the individual turn those desires into binding goals that can potentially be realized and possess a genuine attractiveness. Moreover, deliberating requires a general open-mindedness (as was demonstrated in the Heckhausen and Gollwitzer [1987] study described above) with respect to any available information, as undecided individuals do not know yet in which direction their decision will finally take them.

Once a goal decision has been made, the task of planning is to promote the initiation of goal-directed behavior. This requires committing oneself to when, where, and how to get started. Accordingly, one needs to discover good opportunities and link them to appropriate goal-directed actions, thus creating plans for action. For this purpose, cognitive tuning toward implementation-related issues should be beneficial. Feasibility-related and desirability-related issues should no longer matter, and, if forced on the individual, they are avoided by distorting the relevant information in support of the goal decision made: the person sees the feasibility of the chosen goal in an overly optimistic way, and views the desirability of the chosen goal in a partial manner (i.e., pros exceed cons). Finally, processing all of the available information in an open-minded manner should be dysfunctional, as it might derail the individual from the chosen course of action. Accordingly, a reduced open-mindedness (closed-mindedness) favouring the selective processing of information in support of the chosen goal is to be expected.

Given these different features of the deliberative and implemental mindsets, one should not forget that the two different mindsets also possess many similar attributes. For instance, the mindset theory of action phases assumes that both deliberative and

implemental mindsets become more pronounced as a person gets more involved with deliberating between potential goals and with planning chosen [p. 530 ↓] goals, respectively. Moreover, neither mindset should immediately vanish when the task activity that produced it is ended; instead, the mindset should show a moment of inertia. This implies that the cognitive orientations associated with the deliberative and implemental mindsets can be detected in then-effects on performing temporally subsequent tasks of a different nature. These ideas regarding the similarity between the deliberative and implemental mindset have been used to develop a research program aimed at testing the proposed different cognitive features of the deliberative and implemental mindsets.

In this research, the following method of inducing the deliberative and implemental mindsets turned out to be most effective: experimental participants are asked either to extensively deliberate an unresolved personal problem to be named by the participants (who indicate problems such as, "Should I move to another city or not?," "Should I change my major?," "Should I buy a new car?," or "Should I get involved with somebody?") or to plan the implementation of a chosen goal indicated by the participants (projects such as, "I will move to another city," "I will change my major," etc., are named). These requests create a deliberative and an implemental mindset, respectively. To intensify these mindsets, deliberating participants are asked to list the short-term and long-term pros and cons of making and not making a decision, in order to get heavily involved with deliberating. Planning participants, on the other hand, are asked to list the five most important steps of implementing the chosen goal, and then to specify when, where, and how they intend to execute each step, all of which serves the purpose of creating an intensive involvement with planning. Thereafter, both the deliberating and the planning participants are asked to perform presumably unrelated tasks (usually presented by a different experimenter in a different situational context), which are designed to measure the very cognitive features hypothesized to differ between the deliberative and implemental mindsets. This procedure of inducing the deliberative and implemental mindsets in one situational context and assessing their cognitive and behavioral consequences in a different setting, has been referred to as procedural priming or mindset priming (Bargh and Chartrand, 2000), as research participants commonly stay unaware of the mindset effects they evidence.

Deliberative versus Implemental Mindsets and Cognitive Tuning

The hypothesis that the deliberative mindset creates cognitive tuning toward information relevant to making goal decisions (information on feasibility and desirability), whereas the implemental mindset tunes a person's cognitions to implementation-related information (information on where, when, and how to act), was tested most critically by Gollwitzer et al. (1990). Participants were placed into either a deliberative or an implemental mindset by having them deliberate on unresolved personal problems or plan chosen goal projects, respectively (the standard procedure described above was used). In a presumably unrelated second part of the experiment, participants were presented with the first few lines of a number of novel fairy tales and were instructed to complete each tale. Even though participants were allowed to continue the stories in any way they liked, deliberating participants had the protagonists of the tales reflect on reasons for choosing or not choosing certain action goals to a greater degree than planning participants did. Thoughts about how to accomplish a chosen goal, however, were more frequently attributed to the protagonists by planning participants than by deliberating participants.

Focusing on the processing of mindset-congruent information, Gollwitzer et al. (1990, Study 2) conducted an experiment in which participants had to recall the presented deliberative and implemental thoughts of others. Participants were placed into either a deliberative or an implemental mindset by [p. 531 ↓] having them reflect the choice of one of two tests (i.e., decide between two different creativity tests) or plan to perform a chosen test. While participants were involved in deliberating or planning, slides were presented that depicted different persons mulling over personal decisions. For example, a depicted elderly lady was thinking of the pros (i.e., "It would be good because ...") and cons (i.e., "It would be bad because ...") of having her grandchildren spend their summer vacation at her home. For each of these slides, next to the pros and cons of making a decision, potential plans of implementation were also presented. These specified how the person would get started with the particular goal-directed actions (i.e., "If I decide to do it, then I will first ... and then ...!"; "If I decide to do it, then I won't...").

before ...!"). A cued-recall test of this information was given following a distractor task; it provided participants with the pictures of the persons they had viewed and the stems of the sentences (as above) describing their thoughts. The deliberating participants, who had to view the slides and to recall the information depicted on the slides prior to making a decision about the two types of creativity tests, recalled pros and cons better than they recalled information on the when, where, and how of implementation. The recall performance of the planning participants, who had received and recalled the information after a decision on the creativity tests had been made, showed the reverse pattern.

All of these findings corroborate the cognitive-tuning hypothesis. But how do these differential recall performances observed in the last study (Gollwitzer et al., 1990, Study 2) come about? If one assumes that individuals' retrieval attempts necessitate constructing descriptions of what they are trying to retrieve (Norman and Bobrow, 1979), it seems possible that mindsets provide perspectives (Bobrow and Winograd, 1977) that allow the easy construction of specific descriptions. The deliberative mindset for instance should favour descriptions phrased in terms of pros and cons, benefits and costs, and so forth. In other words, the deliberative mindset supports the ready construction of descriptions that specify desirability-related information, whereas the implemental mindset supports the construction of descriptions that specify implementation-related information. As Norman and Bobrow (1979) point out, quick construction of specific descriptions at the time of retrieval facilitate further successful retrieval. Norman and Bobrow also assume that whenever the description of the information sought matches the elaboration of the information at the time of encoding, recall performance is particularly enhanced. It seems possible, then, that deliberative and implemental mindsets favor congruent recall through both congruent elaboration at the time of encoding and ready construction of congruent descriptions at the time of retrieval.

Deliberative versus Implemental Mindsets and Biased Inferences

Deliberative and implemental mindsets are also postulated to differentially affect the way in which feasibility-related and desirability-related information is handled. In a deliberative mindset, information related to desirability should be analyzed impartially; in an implemental mindset, an analysis partial to the chosen goal is expected. Also, feasibility-related information is expected to be analyzed rather accurately in a deliberative mindset, whereas optimistic inferences that overestimate the actual feasibility of the chosen goal are expected in an implemental mindset.

Desirability-Related Information

With respect to testing the postulated impartial versus partial analysis of desirability-related information, a first study (reported by Taylor and Gollwitzer, 1995, Study 3) was conducted by asking participants to name either potential goals or chosen goals and [p. 532 ↓] subsequently attempt to achieve clarity on the question of whether they should make an affirmative decision or had made the correct decision, respectively. Whereas the predecisional participants reported on positive and negative consequences with the same frequency, postdecisional participants failed to do so. The latter reported about five times more thoughts about pros than about cons, indicating a strong partiality in favor of the chosen goal in postdecisional participants.

Evidence for differences between the deliberative and implemental mindset in processing pros and cons is also provided by Harmon-Jones and Harmon-Jones (2002, Study 2). They tested the effects of mindsets on the postdecisional spreading of alternatives, a classic cognitive dissonance paradigm (Brehm and Cohen, 1962). Using this paradigm, dissonance researchers have found that after making a choice between two options, the chosen option becomes evaluated more positively whereas the nonchosen option becomes evaluated more negatively. Harmon-Jones and Harmon-Jones found that the implemental mindset increased postdecisional spreading of alternatives, whereas the deliberative mindset reduced it.

There is an important set of studies by Gagné and Lydon (2001a) suggesting that deliberation only then leads to an impartial analysis of pros and cons when deliberation is linked to the predecisional action phase. Deliberation over goal decisions that have already been made can initiate defensive processing of information that leads to even greater biasing. In one study, they asked participants involved in romantic relationships to deliberate a relationship or a nonrelationship goal decision. They found that when asked to rate how their partner compared with the average, those individuals asked to deliberate over a relationship goal decision gave much higher ratings than those who were asked to deliberate over a nonrelationship goal decision. Of interest, these ratings were also higher than those of implemental participants who had been planning the implementation of a relationship goal. Gagné and Lydon (2001a) argue that the deliberation of a relationship goal may have been perceived as threatening, resulting in greater enhancement of the partner's attributes. In a second study, they measured the commitment participants had to their relationship and found that high-commitment but not low-commitment participants defended against the threat of a deliberative mindset by increasing their positive views of their partner. This pattern of findings indeed supports the assumption that deliberation may have threatened the participants' perceived ability to attain the goal of maintaining the relationship. In response, these individuals reasserted their commitment to the relationship by boosting the ratings of their partner.

Feasibility-Related Information

The hypothesized accurate analysis of feasibility-related information in the deliberative mindset, and the expected overly optimistic assessment in the implemental mindset, were observed in experiments by Gollwitzer and Kinney (1989) using the contingency-learning task designed by Alloy and Abramson (1979). In this task, participants are asked to determine to what degree they can influence the onset of a target light (outcome) by choosing to press or not to press a button (alternative actions). Participants commonly go through a series of trials (at least 40); the start of each trial is indicated by a warning light. By observing whether or not the target light comes on after they have pressed or not pressed the button, participants estimate how much control they have over the target light onset. The experimenter varies the actual control

by manipulating the frequency of the target light onset associated with each of the two action alternatives (pressing or not pressing). The smaller the difference between these two frequencies, the less objective control participants have over the target light onset.

Nondepressed individuals commonly claim to possess control over target light onset that is noncontingent on their actions, [p. 533 ↓] whenever the target light onset occurs frequently (e.g., in the “75/75” problem, where the target light comes on in 75 percent of pressing and 75 percent of nonpressing responses; see Alloy and Abramson, 1979). Gollwitzer and Kinney (1989, Study 2) asked deliberating, planning, and control participants to work on a contingency problem that presented frequent and noncontingent target light onset (i.e., the 75/75 problem). Participants were given the instruction to discover how to produce the target light onset. A set of 40 trials was offered, and participants were then asked to judge how much control they could exert over the target light onset.

Deliberating participants showed the most accurate judgment of control; their judgments of control were lower than those of either the control group or the planning group. The planning participants' judgments of control tended to be even higher than those of the control participants. The mindsets were created via the standard procedure described above. A mindset interpretation of these findings is supported by the additional observation that deliberating participants' judgments of control correlated negatively with the personal importance of the unresolved personal problems these participants were mulling over. Apparently, the more involved participants were in deliberating, the more realistic their subsequent judgments of control. A parallel finding was observed for planning participants, whose judgments of control were positively related to the participants' anticipated frustration in case they should fail to implement their chosen goals.

When Shelly Taylor asked me to collaborate on studies testing whether deliberative and implemental mindsets might even differentially affect the perceived controllability of events in everyday life, I was happy to assent. We observed (Taylor and Gollwitzer, 1995, Study 1) that the deliberative and implemental mindset indeed manage to affect people's judgments of the controllability of everyday risks (e.g., the risks involved being in an automobile accident, becoming divorced, becoming depressed, developing a drinking problem, and being mugged). Participants were college students who had to

judge these risks for themselves and for the average college student. Mindsets were induced via the standard procedure just before participants had to judge the named risks. Even though all participants perceived themselves as less vulnerable to these risks than the average college student, deliberating participants did this to a lesser degree than planning participants. This more pronounced illusion of invulnerability in the implemental mindset than in the deliberative mindset held no matter whether the critical events to be considered were of a more or less controllable (e.g., developing an addiction to prescription drugs, having a drinking problem) versus uncontrollable kind (e.g., developing diabetes, losing a partner to an early death). For both types of events, planning participants reported a higher invulnerability as compared to the average college student than deliberating participants did. The fact that deliberative and implemental mindsets even managed to modify the perceived vulnerability of rather uncontrollable events attests again to their enormous influence on the analysis of feasibility-related information.

We reasoned that assessing the feasibility of potential goals in the predecisional action phase not only requires that people accurately assess whether their actions could effectively control desired outcomes, they also need to know whether they are in the position to perform these instrumental actions. To answer this question, they have to assess correctly whether they possess the relevant aptitudes and skills. This implies that people in a deliberative mindset should show a relatively accurate evaluation of their personal attributes. Accordingly, we (Taylor and Gollwitzer, 1995, Study 2) also asked deliberating and planning participants to rate themselves on 21 qualities and skills (e.g., cheerfulness, athletic ability, writing ability, popularity, artistic ability) in comparison with the average college student of the same [p. 534 ↓] age and gender. Even though all participants perceived themselves as more capable than the average college student, planning participants did so to a higher degree than deliberating participants. More recent research by Bayer and Gollwitzer (2005) suggests, however, that deliberative and implemental mindset effects on perceived ability may be moderated by people's relevant original low versus high self-views. The deliberative mindset helped in particular people with originally high self-views to arrive at modest ability appraisals (that foster the setting of realistic goals); and it was again in particular people with originally high self-views that were helped by the implemental

mindset to arrive at optimistic ability appraisal (that foster the attainment of chosen goals).

A study by Puca (2001) also speaks to the biased analysis of feasibility-related information in the implemental as compared to the deliberative mindset. She studied realism versus optimism in terms of choosing test materials of different difficulty (Study 1) and predicting their own future task performance (Study 2). Deliberative participants preferred tasks of medium difficulty, whereas implemental participants opted for the too difficult tasks; also, implemental mindset participants overestimated their probability of success more than deliberative participants. Moreover, deliberative participants referred more than implemental participants to their past performance when selecting levels of difficulty or predicting future performance. Finally, when Gagné and Lydon (2001b) moved this biased inferences research to the real world by studying the issue of relationship predictions they found that individuals with a deliberative mindset were more accurate in their forecasts of survival of their romantic relationships than individuals with an implemental mindset. This effect was even more pronounced for long-term than for short-term relationship survival. Of most interest, participants with a deliberative mindset did not achieve this heightened accuracy by simply taking a pessimistic attitude.

Deliberative versus Implemental Mindsets and Open-Mindedness

Beyond differences in cognitive tuning and biased inferences, deliberative and implemental mindsets should also differ in openness to information. Task analysis of the demands of making a goal decision suggests that deliberative mindsets should be associated with enhanced receptivity to all sources and types of information. To make good decisions, one should be open to any available information that might potentially inform one's decision-making. One should be careful not to dismiss information prematurely as it may ultimately be useful or helpful in making good goal decisions. Implemental mindsets, in contrast, should be associated with more selective information processing. Once a goal is set, successful goal implementation requires more particular filtering of information, selectively processing goal-relevant stimuli while ignoring goal-

irrelevant stimuli (e.g., Gollwitzer, 1990; Kuhl, 1984). For these reasons, the deliberative mindset should be associated with greater openness to information incidental to one's goals.

The early study by Heckhausen and Gollwitzer (1987, Study 2) reported above bears some relevance to this hypothesis. Participants were interrupted either while they were deliberating a choice between two different creativity tests (deliberative mindset), or just after having chosen one of them (implemental mindset), and verbally presented with lists of five to seven one-syllable nouns (e.g., house, art, and tree). Immediately after each list had been presented, participants had to recall the words in order. Participants' performance in this task was used to compute their working memory span (i.e., noun span), and results indicated that deliberative mindset participants evidenced a broader span (about half a word more) than implemental mindset participants.

The superior noun span by deliberative as compared with implemental mindset participants, however, only suggests that [p. 535 ↓] deliberative mindset participants are more capable of storing information (i.e., they have a broader working memory span). Although broadened working memory suggests an enhanced capacity to process information, it does not directly address the hypothesis that deliberative mindsets, as compared with implemental mindsets, are associated with heightened processing of information that is incidental to one's goals. The information in the word lists used in the Heckhausen and Gollwitzer (1987) study cannot be considered incidental. Participants were explicitly asked to correctly reproduce as many words as possible of each presented word list. Moreover, broader working memory span by itself does not necessarily lead to more or less selective processing of incidental information.

Accordingly, Fujita et al. (2007) attempted a more critical test of the hypothesis that there are differences between the deliberative and implemental mindset in the selective processing of incidental information. In three experiments, participants in deliberative and implemental mindsets performed a primary task (*d2-concentration test*; Brickenkamp, 1981) while randomly presented incidental, unavoidable words. A subsequent unexpected recognition memory test assessed selective processing of these incidental words. In Study 1, they observed that participants in a deliberative mindset took less time than those in an implemental mindset to recognize whether or not they had previously been exposed to incidental words presented in a concentration

task. In Studies 2 and 3, participants in a deliberative mindset had higher recognition accuracy of these words as compared with those in an implemental mindset. The results from all three studies indicate that deliberative individuals more easily accessed memory traces of information incidental to the ongoing task than implemental individuals. This occurred even when the mindsets induced were unrelated to the performance task that measured the cognitive differences (Studies 2 and 3). This “carryover” effect of mindsets suggests that whereas the implemental mindset is more selective, the deliberative mindset is more open-minded to incidental information available in one's immediate environment. Study 3 also allowed clarifying whether the effect of mindset on selective information processing was due to enhanced open-mindedness in the deliberative mindset, enhanced closed-mindedness in the implemental mindset, or both. Results from this study suggested that the change in selective processing as a function of mindset is attributable to less selective filtering of incidental information in the deliberative mindset, as opposed to greater selective filtering in the implemental mindset.

Finally, the three studies taken together indicate that changes in selective processing as a function of mindset occur preconsciously. Preconscious cognitive processes are those that are initiated and do operate outside of conscious intent (Bargh, 1994). Researchers have argued that reactions to stimuli that require a response within 300 ms are not consciously controlled (e.g., Bargh and Chartrand, 2000; Greenwald and Banaji, 1995). Participants in all three studies were presented with incidental stimuli for only 300 ms prior to the primary performance task materials. Apparently, the act of deliberating a goal decision produces dramatic changes in the cognitive processing of information even when individuals do not intend such changes.

Deliberative versus Implemental Mindsets and Behavior

Gollwitzer and Bayer (1999) pointed out that mindsets have been analyzed primarily in terms of their cognitive features, whereby the effects of these features on the control of behavior were ignored. As an exception, they reported a study by Pösl (1994), who found that participants in the implemental mindset were faster to initiate goal-

directed behavior than those in the deliberative mindset. This was particularly true when participants experienced a behavioral conflict (i.e., whether [p. 536 ↓] they had a choice to perform Behavior A or B, or only one of these), suggesting that the implemental mindset's closed mindedness allows participants to stay on track even in the case of behavioral conflict.

There is also evidence that the implemental mindset generates greater persistence in goal-directed behavior. Brandstätter and Frank (2002) found that participants in the implemental mindset persisted longer at an unsolvable puzzle task (Study 1) and at a self-paced computer task (Study 2). Similar to the findings of Pösl (1994), the impact of the implemental mindset on persistence was present in particular in situations of behavioral conflict. When both the perceived feasibility and desirability of the tasks were either uniformly high or low, persistence on the persistence tasks did not differ by mindset. However, when the perceived feasibility and desirability of the tasks were in opposite directions (i.e., one was high whereas the other was low) the implemental mindset participants persisted longer than did the deliberative mindset participants. Interestingly, the persistence in the implemental mindset was not executed insensitively or in a blind fashion. Brandstätter and Frank (2002, Study 3) obtained evidence that when a task is perceived as impossible or when persistence is not beneficial, individuals in the implemental mindset disengage much more quickly than individuals in the deliberative mindset.

Finally, Armor and Taylor (2003) report that an implemental mindset facilitates task performance (i.e., a scavenger hunt to be performed on campus) as compared with a deliberative mindset, and that this effect is mediated by the cognitive features of the implemental mindset (i.e., enhanced self-efficacy, optimistic outcome expectations, perceiving the task as easy). Such a study had been missing so far. Note that the Gollwitzer and Kinney (1989) study predicted cognitive changes (i.e., strong illusions of control) as a consequence of the implemental as compared with the deliberative mindset, because such illusions should benefit acting on the goal. However, this inference had not been tested within one and the same study as was done by Armor and Taylor.

Recent research by Henderson et al. (2008) suggests that the beneficial effects of the implemental mindset on task performance might also be mediated by changes

in respective attitude strength. The unambiguous, evaluative polarized or one-sided assessment of information in support of the chosen goal in the implemental mindset should foster the strength of the attitude toward the goal. Given the carryover properties of mindsets, Henderson et al. hypothesized that as people adopt an implemental mindset, they should experience an increase in attitude strength toward objects even if these are unrelated to their current goal pursuit. In a series of experiments they found support for this hypothesis. They observed that implemental mindset participants more than deliberative mindset participants adopt an extreme position towards an issue that is irrelevant to their goal concern. Moreover, implemental mindset participants evidenced lower levels of ambivalence toward a variety of unrelated objects than deliberative and neutral mindset participants. Implemental mindset participants were also characterized by more accessible evaluations of unrelated objects than deliberative and neutral mindset participants. And, importantly, implemental mindset participants showed a greater correspondence between their attitude and behavior than neutral mindset participants.

Finally, in order to investigate the process that is hypothesized to underlie the effects of an implemental mindset on attitude strength (i.e., one-sided focus on pros), Henderson et al. had implemental mindset participants either focus on the pros only or on both the pros and cons behind their decision. Critically, it was only the evaluative one-sided analysis of information that fostered attitude strength. As the authors had taken two groups who made a decision on how to act and only varied their evaluative focus (i.e., one group engaged in a one-sided evaluation of their [p. 537 ↓] decision, whereas the other group engaged in a two-sided evaluation of their decision), the act of deciding itself does not seem to be sufficient to increase attitude strength. Otherwise, those individuals who made a decision on how to act and who analyzed both sides of their decision would have evidenced the same level of attitude strength as those who analyzed only one side of their decision.

Summary

Under the assumption that the course of goal pursuit presents itself to the individual as a series of consecutive tasks that need to be solved in order to promote goal attainment, the concept of mindset has been introduced. Mindset theory of action

phases argues that becoming involved in these tasks leads to characteristic cognitive orientations (mindsets) that are beneficial for solving these tasks effectively, and the features of the cognitive orientations associated with the tasks of choosing between potential action goals (the deliberative mindset) and preparing the implementation of chosen goals (the implemental mindset) are spelled out. Various experiments tested the postulated characteristics of the deliberative and implemental mindsets. This research shows that the deliberative mindset is characterized by cognitive tuning toward desirability-related and feasibility-related thoughts and information, by an accurate analysis of feasibility-related information and an impartial analysis of desirability-related information, and, finally, by a heightened general receptivity to available information. The implemental mindset, on the other hand, is characterized by cognitive tuning toward implemental thoughts and information, by an overly optimistic analysis of feasibility-related information and a partial analysis of desirability-related information, and, finally, by a comparatively reduced receptivity (closed-mindedness) to available information.

Implications for other Theories and Conceptual Debates in Social Psychology

Mindset theory of action phases has received much attention in other theories and conceptual debates in social psychology. These pertain to the optimism versus realism debate and to dual process theories. How the mindset theory of action phases can help to clarify issues in these areas of research will be discussed next.

Optimism versus Realism

The results of mindset research on the processing of feasibility-related information pertain to the illusionary optimism versus realism controversy triggered by Taylor and Brown's (1988) article on positive illusions. Taylor and Brown proposed that mentally healthy people are not characterized by accurate assessments of their personal qualities, realistic estimates of personal control, and a realistic outlook on the future; instead, they maintain overly positive, self-aggrandizing perspectives of the self, the world, and the future. More specifically, mentally healthy people are

said to be characterized by unrealistically positive self-perceptions, an illusion of a high degree of personal control, and unrealistic optimism about the future. Instead of being maladaptive, these positively distorted perceptions foster the criteria normally associated with mental health: positive regard, the ability to care for and about other people, and the ability to manage stress effectively (Taylor and Brown, 1988). Despite empirical support for the model (Taylor and Armor, 1996; Taylor et al., 2000), this portrait of the healthy person raises a disturbing question: if healthy people's perceptions are marked by positive bias, how do they effectively identify and make use of negative feedback they may encounter? If people are capable of explaining away, compartmentalizing, or otherwise dismissing or minimizing negative feedback, as Taylor and Brown [p. 538 ↓] (1988) suggested, these self-serving illusions that bolster self-esteem and produce a positive mood in the short run may ultimately set people up for long-term disappointment and failure as they fail to incorporate negative feedback into their goal setting and planning (Colvin and Block, 1994; Weinstein, 1984).

The mindset research on illusion of control offers the following insights to the debate about positive illusions versus realism. First, neither realism nor positive illusions seem adaptive in general to a person's psychological functioning. Realistic thinking seems functional when it comes to making goal decisions, whereas positive illusions seem functional when the chosen goals are to be implemented. Second, people can easily open the window to realism provided by the deliberative mindset. People do not have to go through the effortful mental exercises we have induced in our experiments to create a deliberative mindset; simply trying to achieve clarity in regard to an unresolved personal problem will trigger an intensive deliberation of pros and cons (Taylor and Gollwitzer, 1995, Study 3). Third, postdecisional individuals who plan the implementation of a chosen goal seem to be protected from an accurate analysis of feasibility-related information and thus can benefit from illusionary optimism that makes them strive harder to reach their goals, especially in the face of hindrances and barriers. It appears, then, that the individual's cognitive apparatus readily adjusts to the various demands of the control of action: choosing between action goals leads to realism, and implementing chosen goals leads to positive illusions.

Dual Process Theories

The ideas and research originating within the framework of the notion of deliberative and implemental mindsets seem to constitute a dual process theory in the realm of goal pursuit. The approach taken is to juxtapose a cognitive orientation that is functional to choosing goals with a cognitive orientation that is functional to the implementation of chosen goals. In other words, the ideal information-processing styles for solving two different tasks that serve one end (i.e., the effective control of action) are analyzed in contrast to each other. This is different from those dual process models that compare two different styles of information processing in the service of one and the same task, such as perceiving another person (Bargh, 1984; Brewer, 1988; Fiske and Neuberg, 1990), making attributions (Gilbert, 1989; Gilbert and Malone, 1995), or forming attitudes (Chaiken et al., 1989; Fazio, 1990). The approach taken in those models is to analyze how the two forms of information processing delineated differ in meeting the task at hand as is also true for those dual process lines of research that explicitly adopt the mindset notion for conceptual and methodological reasons, such as research on the counter-factual mindset (e.g., Galinsky and Moskowitz, 2000; Wong et al., 2009) and near versus distal mental construal (e.g., Liberman and Trope, 2008; Freitas et al., 2004). Counterfactual mindset research induces a counterfactual mindset by having participants read one and the same behavioral episode experienced by another person that is known to trigger counterfactual musings (“If only he had done ...,” or “What if he had done ...”), and then checks whether reading as compared to not reading this episode improves performance in subsequent classic problem solving, creativity, or negotiation tasks. Mental construal research, on the other hand, studies how inducing psychologically distant versus near construals of certain events affects the perception, categorization, judgmental inferences, evaluations, and behaviors with respect to these events as well as related events. Note that the attempted gain in knowledge of these two approaches refers to how a given cognitive orientation (counterfactual, near versus distal mental construal) affects respective cognitive, affective, and behavioral performances. With respect to the mindset theory of action phases, [p. 539 ↓] however, the attempted gain of knowledge refers to the typical characteristics of deliberating versus planning in terms of the underlying cognitive procedures.

However, there are some similarities between the deliberative versus implemental mindset distinction and those dual process notions that are construed as stage theories (e.g., Gilbert and Malone, 1995). People in everyday life should experience deliberative mindsets prior to implemental mindsets as people mostly prefer to make plans on how to achieve a goal only after they have made a binding goal choice. In this temporal sense therefore the deliberative versus implemental mindset model qualifies as a stage model. This is not true, however, with respect to the quality of cognitive processes associated with the two mindsets. For instance, in the two-step model of the attribution process – a stage model suggested by Gilbert (1989; Gilbert and Malone, 1995) – the first step is simple and automatic (i.e., a quick personal attribution), whereas the second step requires attention, thought, and effort (i.e., adjusting that inference to account for situational influences). The notion of deliberative versus implemental mindsets, on the other hand, does not assume that the deliberative mindset is associated with more rudimentary cognitive processes than the implemental mindset (or vice versa). In both the deliberative mindset and the implemental mindset, highly complex cognitive procedures are activated that determine the individual's cognitive and behavioral functioning. Moreover, in both deliberative and implemental mindsets these procedures can, but do not need to, reach consciousness to unfold their effects, and their effects can, but do not have to, be detected by the individual (e.g., the illusion of control in the implemental mindset).

Also, mindset theory sees the deliberative and implemental mindsets as distinct and independent of each other. Whereas it is assumed in some dual process theories (e.g., Chaiken et al., 1989) that the postulated modes of information processing can operate at the same time, the deliberative and implemental mindsets are assumed to preclude each other. This is because the strength with which the cognitive procedures associated with the deliberative mindset are activated is positively related to the degree of involvement with the task of choosing between potential goals, whereas the strength with which the cognitive procedures associated with the implemental mindset are activated is positively related to the degree of involvement with the task of planning the implementation of a chosen goal. Because a person cannot become intensely involved in both of these tasks at one and the same time but only successively, pronounced deliberative and implemental mindsets cannot coexist. They also do not affect each other in the sense that a preceding strong deliberative mindset makes for a strong

succeeding implemental mindset; it all depends on how intensely people become involved with solving the task of choosing between potential goals and with planning the implementation of a chosen goal, respectively.

Summary

The deliberative versus implemental mindset distinction had an impact on various theoretical discussions in social psychology. First, with respect to the realism versus optimism discussion, it made clear that people are capable of flexibly adopting the type of orientation that is demanded by the task at hand. When decisions between goals are to be made (e.g., whether to go to college at home or abroad), getting involved with reflections on which goal to choose spurs realism that in turn allows making the more appropriate (feasible) choice. And if a chosen goal (e.g., studying abroad) is to be implemented, getting involved with planning out the course of goal realization spurs optimism that promotes the necessary persistence for goal attainment. Second, with respect to classic dual process theories, the mindset theory of action phases points out that modes of information processing may not only be studied in [p. 540 ↓] terms of what determines taking one or the other mode (e.g., heuristic versus systematic information processing is found to be affected by time pressure, likeability of the source, strong versus weak arguments), and what effects taking one or the other mode has on grasping the information at hand (e.g., leads to more or less respective attitude change). Mindset theory of action phases suggests that the mere involvement with one or the other type of reasoning task (deliberating the choice of potential goals versus planning the implementation of a chosen goal) already activates different task-facilitating cognitive procedures that affect the processing of both task-relevant and task-irrelevant information in a unique way.

The Applied Impact of the Mindset Theory of Action Phases

Mindset theory of action phases had set aside the question of what makes for good planning of goal implementation. So when I was invited to create a research unit on

“Intention and Action” at the Max Planck Institute for Psychological Research at Munich in 1989, this question became our central concern and has stayed with me ever since. As we had induced the implemental mindset by asking research participants to list a series of steps toward goal attainment and then specify for each individual step exactly when, where, and how one wants to realize it, we wondered whether people could facilitate goal striving by planning out goal-directed action in this fashion. We referred to this type of planning as forming *implementation intentions* (Gollwitzer, 1993, 1999). Whereas goals (or goal intentions) merely specify desired end states (“I want to achieve goal X!”), implementation intentions in the format of “If situation Y arises, then I will initiate behavior Z!” additionally specify when, where, and how a person intends to strive for the goal. Implementation intentions thus delegate control over the initiation of the intended goal-directed behavior to a specified situational cue by creating a strong mental link between this cue and a goal-directed response. For example, a person who has chosen the goal to eat more healthily can form the implementation intention, “When I’m at my favorite restaurant and the waiter asks me for my order, then I’ll request a vegetarian meal!” The mental links created by implementation intentions were assumed (Gollwitzer, 1999) to facilitate goal attainment on the basis of psychological processes that relate to both the anticipated situation (i.e., enhanced activation of the mental representation of the situation specified in the if-part of the plan) and the intended behavior (i.e., automatic initiation of the response specified in the then-part of the plan once the critical situation is encountered).

I was fortunate to get to know Paschal Sheeran in the mid 1990s. His students and colleagues in England as well as mine in Germany set out to test the claimed positive effects of implementation intentions on goal attainment as well as the assumed underlying processes. Because forming an implementation intention implies the selection of a critical future situation, the mental representation of this situation becomes highly activated and hence more accessible. This heightened accessibility of the “if” part of the plan has been observed in numerous studies; it helps people to easily recall the specified situation and it leads to swift allocation of attention when the situation arises. Various studies also observed that the initiation of the goal-directed response specified in the then-component of an implementation intention exhibits features of automaticity; if-then planners are found to act quickly in the face of the

critical situation, they deal effectively with cognitive demands, and do not need to consciously intend to act in the critical moment.

A meta-analysis involving over 8,000 participants in 94 independent studies revealed a medium-to-large effect size of implementation intentions on goal achievement (Gollwitzer and Sheeran, 2006). [p. 541 ↓] Importantly, implementation intentions were found to benefit goal-directed responses no matter whether these were cognitive, emotional, or behavioral in nature, and this held true in all kinds of goal domains (e.g., consumer, environmental, antiracist, prosocial, academic, and health). Apparently, implementation intentions help people to better cope with the major problems of goal striving: getting started, staying on track, calling a halt to a futile goal striving, and not over-extending oneself. Importantly, implementation intentions still show their beneficial effects when the going gets tough (Gollwitzer et al., 2010; Gollwitzer & Oettingen, 2011); that is, when goal striving is limited by conditions that are very resistant to change by self-regulatory strategies (e.g., a low level of competence, a fierce competitor, or strong competing habitual responses).

A recent fMRI study conducted by Gilbert et al. (2009) provides an answer to the puzzling power of implementation intentions. Brain activity in the lateral area 10 was observed to move toward the medial area 10 when participants switched from performing an executive-function task by the guidance of a goal intention to performing the very same type of task by the guidance of an implementation intention. On the basis of an extensive meta-analysis on various executive-function tasks it is known that lateral and medial area 10 are implicated in top-down and bottom-up action control, respectively (Burgess et al., 2005). Apparently, implementation intentions induce a switch in action control from top-down to bottom-up control of action. This explains why even habitual responses can be broken by implementation intentions, and why special populations that are known to suffer from ineffective conscious control of their thoughts, feelings, and actions are also found to benefit from forming implementation intentions (e.g., heroin addicts during withdrawal, schizophrenic patients, children with attention deficit hyperactivity disorder [ADHD]).

But will the discovery of an effective self-regulation strategy of goal striving and the in-depth analysis of how it works make people use it to solve personal, interpersonal, and even societal problems? Simply disseminating the good news of the existence

of a powerful strategy for goal implementation may not suffice. It needs the second step of developing an intervention that facilitates the acquisition of this strategy so that people can use it on their own in everyday life. In other words, one needs to develop an intervention that effectively teaches the forming of implementation intentions as a meta-cognitive strategy.

To come up with such an effective intervention, Gabriele Oettingen and I first considered all of the studies that had analyzed potential moderators of implementation effects (see Gollwitzer and Sheeran, 2006). Importantly, implementation intentions only then unfold their beneficial effects if strong goal commitments are in place. Moreover, implementation intention effects seem to be stronger when people put exactly those cues in the if-part of an implementation intention which they personally consider to be most critical (Adriaanse et al., 2009). Accordingly, we looked for a procedure that established these prerequisites, and we found it in *mental contrasting* as this self-regulation strategy is also known to motivate if-then planning.

What is mental contrasting and how does it work? When people mentally contrast (Oettingen et al., 2001), they first imagine a desired future (e.g., to improve one's health behavior), and then reflect on the present reality that stands in the way of reaching this desired future (e.g., feeling the urge to give in to a temptation). Thereby, mental contrasting turns desired futures that are perceived as feasible into strong goal commitments. The beneficial effects of mental contrasting evince in various domains (achievement, interpersonal, and health), for cognitive (e.g., making plans), affective (e.g., feelings of anticipated disappointment in case of failure), motivational (e.g., feelings of energization, systolic blood pressure), and behavioral indicators of goal commitment (e.g., invested effort and [p. 542 ↓] actual achievement such as obtained course grades), directly after or weeks later (summary by Oettingen and Stephens, 2009). Mental contrasting facilitates strong goal commitments by changing implicit cognition. The increased strength of association between future and reality induced by mental contrasting has been found to mediate the cognitive, emotional, and behavioral indicators of strong goal commitment, measured by self-report and other-rated performance (Kappes and Oettingen, 2011).

We therefore integrated the forming of implementation intentions with mental contrasting into one self-regulation strategy called Mental Contrasting with

Implementation Intentions (MCII; Oettingen and Gollwitzer, 2010). When taught as a meta-cognitive strategy that is then applied by participants in everyday life, it supports behavior change more than mental contrasting and implementation intentions used in isolation; moreover, MCII effects hold for various life domains (academic achievement, interpersonal relations, romantic satisfaction, and health, eating and regular exercise), and over time periods of up to two years (Adriaanse et al., 2010; Christiansen et al., 2010; Stadler et al., 2009, 2010). When engaging in MCII, participants first go through a mental contrasting exercise to create strong goal commitment and to identify the obstacles that truly stand in the way of goal attainment. Implementation intentions (if-then plans) are then formed to help translate the goal commitment into instrumental behavior by putting the obstacles identified in mental contrasting as critical cues in the if-part of the plan and linking it to an instrumental coping response specified in the then-part.

Conclusion

There are times when people need to make decisions, and there are times when the decisions made have to be implemented. From the perspective of effective action control in everyday life, then, it seems helpful to activate the respective cognitive procedures that facilitate goal setting and goal implementation when making decisions versus acting on them is at issue. In other words, people should allow for and become involved in deliberative or implemental mindsets, depending on whether a goal decision or the implementation of a chosen goal is called for.

Moreover, there exist powerful strategies of pre- and postdecisional reasoning that are more effective than others. Accordingly, interventions geared at helping people to maximize their goal setting and goal striving should not merely confront people with the tasks of committing to goals and implementing the chosen goals. Rather, they should go one step further and equip people with those goal setting (e.g., mental contrasting) and goal implementation strategies (e.g., forming implementation intentions) that are known to be most effective in promoting appropriate goal commitments and successful goal attainment, respectively. Future research might explore how such strategies are taught best in time- and cost-effective behavior change interventions.

References

- Ach, N. (1935) Analyse des Willens (The analysis of willing) . In E. Abderhalden (ed.) Handbuch der biologischen Arbeitsmethoden , Vol. 6. Berlin: Urban u. Schwarzenberg.
- Adriaanse, M.A., de Ridder, D.T.D. and de Wit, J.B.F. (2009) Finding the critical cue: Implementation intentions to change one's diet work best when tailored to personally relevant reasons for unhealthy eating . *Personality and Social Psychology Bulletin* , 35, 60–71. <http://dx.doi.org/10.1177/0146167208325612>
- Adriaanse, M.A., Oettingen, G., Gollwitzer, P.M., Hennes, E.P., de Ridder, D.T.D. and de Wit, J.B.F. (2010) When planning is not enough: Fighting unhealthy snacking habits by Mental Contrasting with Implementation Intentions (MCII) . *European Journal of Social Psychology* 40, 1277–1293. <http://dx.doi.org/10.1002/ejsp.730>
- Alloy, L.B. and Abramson, L.Y. (1979) Judgement of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General* , 108, 449–485. <http://dx.doi.org/10.1037/0096-3445.108.4.441>
- Armor, D.A. and Taylor, S.E. (2003) The effects of mindset on behavior: Self-regulation in deliberative and implemental frames of mind . *Personality and Social Psychology Bulletin* , 29, 86–95. <http://dx.doi.org/10.1177/0146167202238374>
- Atkinson, J.W. (1957) Motivational determinants of risk-taking behavior . *Psychological Review* , 64, 359–372. <http://dx.doi.org/10.1037/h0043445>
- Bargh, J.A. (1984) Automatic and conscious processing of social information . In R.S. Wyer, Jr., ed. and T.K. Srull (eds), *Handbook of Social Cognition* , 3, 1–43. Hillsdale, NJ: Erlbaum.
- Bargh, J.A. (1994) The Four Horsemen of automaticity: Awareness, efficiency, intention, and control in social cognition . In R.S. Wyer, ed. Jr. and T.K. Srull (eds), *Handbook of Social Cognition* , 2nd Edition, pp. 1–40. Hillsdale, NJ: Erlbaum.

Bargh, J.A. and Chartrand, T.L. (2000) A practical guide to priming and automaticity research . In H. Reis, ed. and C. Judd (eds), *Handbook of Research Methods in Social Psychology* , pp. 253–285. New York: Cambridge University Press.

Bayer, U.C. and Gollwitzer, P.M. (2005) Mindset effects on information search in self-evaluation . *European Journal of Social Psychology* , 35, 313–327. <http://dx.doi.org/10.1002/ejsp.247>

Bobrow, D.G. and Winograd, T. (1977) An overview of KRL: A knowledge representation language . *Cognitive Science* , 1, 3–46. http://dx.doi.org/10.1207/s15516709cog0101_2

Brandstätter, V. and Frank, E. (2002) Effects of deliberative and implemental mindsets on persistence in goal-directed behavior . *Personality and Social Psychology Bulletin* , 28, 1366–1378.

Brehm, J.W. and Cohen, A.R. (1962) Explorations in Cognitive Dissonance . New York: Wiley. <http://dx.doi.org/10.1037/11622-000>

Brewer, M.B. (1988) A dual process model of impression formation . In T.K. Srull, ed. and R.S. Wyer (eds), *Advances in Social Cognition* , 1, 1–36. Hillsdale, NJ: Erlbaum.

Brickenkamp, R. (1981) Test d2 , 7th Edition. Goettingen: Hogrefe.

Burgess, P.W., Simons, J.S., Dumontheil, I. and Gilbert, S.J. (2005) The gateway hypothesis of rostral PFC function . In J. Duncan, ed. , L. Phillips, ed. , and P. McLeod (eds) *Measuring the Mind: Speed Control and Age* , pp. 215–246. Oxford: Oxford University Press.

Chaiken, S., Liberman, A. and Eagly, A.H. (1989) Heuristic and systematic information processing within and beyond the persuasion context . In J.S. Uleman, ed. and J.A. Bargh (eds), *Unintended Thought* , pp. 212–252. New York: Guilford Press.

Christiansen, S., Oettingen, G., Dahme, B. and Klinger, R. (2010) A short goal-pursuit intervention to improve physical capacity: A randomized clinical trial in chronic back pain patients . *Pain* , 149, 444–452. <http://dx.doi.org/10.1016/j.pain.2009.12.015>

Colvin, C.R. and Block, J. (1994) Do positive illusions foster mental health?: An examination of the Taylor and Brown formulation . *Psychological Bulletin* , 116, 3–20.
<http://dx.doi.org/10.1037/0033-2909.116.1.3>

Fazio, R.H. (1990) Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 23, 75–109. San Diego: Academic Press.

Feather, N.T. and Newton, J.W. (1982) Values, expectations, and the prediction of social action: An expectancy-valence analysis . *Motivation and Emotion* , 6, 217–244.
<http://dx.doi.org/10.1007/BF00992246>

Fiske, S.T. and Neuberg, S.L. (1990) A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation . *Advances in Experimental Social Psychology* , 23, 1–73.
<http://dx.doi.org/10.1016/S0065-2601%2808%2960317-2>

Freitas, A.L., Gollwitzer, P.M. and Trope, Y. (2004) The influence of abstract and concrete mindsets on anticipating and guiding others' self-regulatory efforts . *Journal of Experimental Social Psychology* , 40, 739–752. <http://dx.doi.org/10.1016/j.jesp.2004.04.003>

Fujita, K., Gollwitzer, P.M. and Oettingen, G. (2007) Mindsets and preconscious open-mindedness to incidental information . *Journal of Experimental Social Psychology* , 43, 48–61. <http://dx.doi.org/10.1016/j.jesp.2005.12.004>

Gagné, F.M. and Lydon, J.E. (2001a) Mindset and relationship illusions: The moderating effects of domain specificity and relationship commitment . *Personality and Social Psychology Bulletin* , 27, 1144–1155.

Gagné, F.M. and Lydon, J.E. (2001b) Mindset and close relationships: When bias leads to (in)accurate predictions . *Journal of Personality and Social Psychology* , 81, 85–96.

Galinsky, A.D. and Moskowitz, G.B. (2000) Counterfactuals as behavioral primes: Priming the simulation heuristic and consideration of alternatives . *Journal*

of Experimental Social Psychology , 36, 384–409. <http://dx.doi.org/10.1006/jesp.1999.1409>

Gilbert, D.T. (1989) Thinking lightly about others: Automatic components of the social inference process . In J.S. Uleman, ed. and J.A. Bargh (eds), Unintended Thought , pp. 189–211. New York: Guilford Press.

Gilbert, D.T. and Malone, P.S. (1995) The correspondence bias . Psychological Bulletin , 117, 21–38. <http://dx.doi.org/10.1037/0033-2909.117.1.21>

Gilbert, S.J., Gollwitzer, P.M., Cohen, A-L., Oettingen, G. and Burgess, P.W. (2009) Separable brain systems supporting cued versus self-initiated realization of delayed intentions . Journal of Experimental Psychology: Learning, Memory, and Cognition , 35, 905–915. <http://dx.doi.org/10.1037/a0015535>

Gollwitzer, P.M. (1987) Motivationale vs. volitionale Bewußtseinslage . (Motivational vs. volitional mindset.) Habilitationsschrift, Ludwig-Maximilians-Universität München.

Gollwitzer, P.M. (1990) Action phases and mindsets . In E.T. Higgins, ed. and R.M. Sorrentino (eds), Handbook of Motivation and Cognition: Foundations of Social Behavior , 2, 53–92. New York: Guilford Press.

Gollwitzer, P.M. (1991) Abwägen und Planen: Bewußtseinslagen in verschiedenen Handlungsphasen (Deliberating and planning: Mindsets in different action phases) . Goettingen: Hogrefe.

Gollwitzer, P.M. (1993) Goal achievement: The role of intentions . European Review of Social Psychology , 4, 141–185. <http://dx.doi.org/10.1080/14792779343000059>

Gollwitzer, P.M. (1999) Implementation intentions: Strong effects of simple plans . American Psychologist , 54, 493–503. <http://dx.doi.org/10.1037/0003-066X.54.7.493>

Gollwitzer, P.M. and Bayer, U.C. (1999) Deliberative and implemental mindsets in the control of action . In S. Chaiken, ed. and Y. Trope (eds), Dual-process Theories in Social Psychology , pp. 403–422. New York: Guilford Press.

Gollwitzer, P.M., Gawrilow, C. and Oettingen, G. (2010) The power of planning: Effective self-regulation of goal striving . In R. Hassin, ed. , K. Ochsner, ed. , and Y. Trope (eds), *Self-control in Society, Mind, and Brain* , pp. 279–296. New York: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780195391381.003.0015>

Gollwitzer, P.M., Heckhausen, H. and Steller, B. (1990) Deliberative and implemental mindsets: Cognitive tuning toward congruous thoughts and information . *Journal of Personality and Social Psychology* , 59, 1119–1127. <http://dx.doi.org/10.1037/0022-3514.59.6.1119>

Gollwitzer, P.M. and Kinney, R.F. (1989) Effects of deliberative and implemental mindsets on illusion of control . *Journal of Personality and Social Psychology* , 56, 531–542. <http://dx.doi.org/10.1037/0022-3514.56.4.531>

Gollwitzer, P.M. and Kirchhof, O. (1998) The willful pursuit of identity . In J. Heckhausen, ed. and C.S. Dweck (eds), *Life-span Perspectives on Motivation and Control* , pp. 389–423. New York: Cambridge University Press.

Gollwitzer, P.M. and Oettingen, G. (2011) Planning promotes goal striving . In K.D. Vohs, ed. and R.F. Baumeister (eds), *Handbook of Self-regulation: Research, Theory, and Applications* , 2nd Edition, pp. 162–185. New York: Guilford Press.

Gollwitzer, P.M. and Sheeran, P. (2006) Implementation intentions and goal achievement: A meta-analysis of effects and processes . *Advances in Experimental Social Psychology* , 38, 69–119. <http://dx.doi.org/10.1016/S0065-2601%2806%2938002-1>

Gollwitzer, P.M., Sheeran, P., Michalski, V. and Seifert, A.E. (2009) When intentions go public: Does social reality widen the intention-behavior gap? *Psychological Science* , 20, 612–618. <http://dx.doi.org/10.1111/j.1467-9280.2009.02336.x>

Greenwald, A.G. and Banaji, M.R. (1995) Implicit social cognition: Attitudes, self-esteem, and stereotypes . *Psychological Review* , 102, 4–21. <http://dx.doi.org/10.1037/0033-295X.102.1.4>

Harmon-Jones, C., Schmeichel, B.J. and Harmon-Jones, E. (2009) Symbolic self-completion in academia: Evidence from department web pages and email signature files . European Journal of Experimental Social Psychology , 39, 311–316. <http://dx.doi.org/10.1002/ejsp.541>

Harmon-Jones, E. and Harmon-Jones, C. (2002) Testing the action-based model of cognitive dissonance: The effect of action orientation on postdecisional attitudes . Personality and Social Psychology Bulletin , 28, 711–723. <http://dx.doi.org/10.1177/0146167202289001>

Heckhausen, H. (1977) Achievement motivation and its constructs: A cognitive model . Motivation and Emotion , 4, 283–329. <http://dx.doi.org/10.1007/BF00992538>

Heckhausen, H. (1987) Wünschen – Wählen – Wollen . In H. Heckhausen, ed. , P.M. Gollwitzer, ed. , and F.E. Weinert (eds), Jenseits des Rubikon. Der Wille in den Humanwissenschaften (Crossing the Rubicon: The concept of will in the life sciences), pp. 3–9. Berlin: Springer Verlag. http://dx.doi.org/10.1007/978-3-642-71763-5_1

Heckhausen, H. and Gollwitzer, P.M. (1987) Thought contents and cognitive functioning in motivational versus volitional states of mind . Motivation and Emotion , 11, 101–120. <http://dx.doi.org/10.1007/BF00992338>

Henderson, M.D., de Liver, Y. and Gollwitzer, P.M. (2008) The effects of implemental mindset on attitude strength . Journal of Personality and Social Psychology , 94, 394–411. <http://dx.doi.org/10.1037/0022-3514.94.3.396>

Higgins, E.T., ed. and Sorrentino, R.M., ed. (1990) (eds) The Handbook of Motivation and Cognition: Foundations of Social Behavior, Vol. 2 . New York: Guilford Press.

Kappes, A. and Oettingen, G. (2011) From wishes to goals: Mental contrasting connects future and reality . Submitted for publication.

Külpe, O. (1904) Versuche über Abstraktion (Experiments on abstraction). Bericht über den 1 . Kongreß für experimentelle Psychologie , pp. 56–68. Leipzig: Barth.

Kuhl, J. (1984) Volitional aspects of achievement motivation and learned helplessness: Toward a comprehensive theory of action control . In B.A. Maher (ed.), *Progress in Experimental Personality Research* , 13, 99–171. New York: Academic Press.

Ledgerwood, A., Liviatan, I. and Camevale, P.J. (2007) Group-identity completion and the symbolic value of property . *Psychological Science* , 18, 873–878. <http://dx.doi.org/10.1111/j.1467-9280.2007.01994.x>

Lewin, K. (1926) Vorsatz, Wille und Bedürfnis (Intention, will, and need) . *Psychologische Forschung* , 7, 330–385. <http://dx.doi.org/10.1007/BF02424365>

Liberman, N. and Trope, Y. (2008) The psychology of transcending the here and now . *Science* , 322, 1201–1205. <http://dx.doi.org/10.1126/science.1161958>

Marbe, K. (1915) Zur Psychologie des Denkens (The psychology of reasoning) . *Fortschritte der Psychologie und ihre Anwendungen* , 3, 1–42.

Mischel, W. (1974) Processes in delay of gratification . *Advances in Experimental Social Psychology* , 7, 249–292. <http://dx.doi.org/10.1016/S0065-2601%2808%2960039-8>

Mischel, W. and Patterson, C.J. (1978) Effective plans for self-control in children . *Minnesota Symposia on Child Psychology* , 11, 199–230.

Norman, D.A. and Bobrow, D.G. (1979) Descriptions: An intermediate stage in memory retrieval . *Cognitive Psychology* , 11, 107–123. <http://dx.doi.org/10.1016/0010-0285%2879%2990006-9>

Oettingen, G., Pak, H. and Schnetter, K. (2001) Self-regulation of goal-setting: Turning free fantasies about the future into binding goals . *Journal of Personality and Social Psychology* , 80, 736. <http://dx.doi.org/10.1037/0022-3514.80.5.736>

Oettingen, G. and Gollwitzer, P.M. (2010) Strategies of setting and implementing goals: Mental contrasting and implementation intentions . In J.E. Maddux, ed. and J.P. Tangney (eds), *Social Psychological Foundations of Clinical Psychology* pp. 114–135. New York: Guilford Press.

Oettingen, G. and Stephens, E.J. (2009) Fantasies and motivationally intelligent goal setting . In G.B. Moskowitz, ed. and H. Grant (eds), *The Psychology of Goals* , pp. 153–178. New York: Guilford Press.

Orth, J. (1903) Gefühl und Bewußtseinslage: Eine kritisch-experimentelle Studie (Feelings and mindsets: An experimental study) . In T. Ziegler, ed. and T. Ziehen (eds), *Sammlung von Abhandlungen aus dem Gebiet der Pädagogischen Psychologie und Physiologie* , 4, 225–353. Berlin: Verlag von Reuter and Reichard.

Pösl, I. (1994) Wiederaufnahme unterbrochener Handlungen: Effekte der Bewusstseinslagen des Abwägens und Planens (Mindset effects on the resumption of disrupted activities) . Unpublished master's thesis, University of Munich, Germany.

Puca, R.M. (2001) Preferred difficulty and subjective probability in different action phases . *Motivation and Emotion* , 25, 307–326. <http://dx.doi.org/10.1023/A:1014815716476>

Stadler, G., Oettingen, G. and Gollwitzer, P.M. (2009) Physical activity in women. Effects of a self-regulation intervention . *American Journal of Preventive Medicine* , 36, 29–34. <http://dx.doi.org/10.1016/j.amepre.2008.09.021>

Stadler, G., Oettingen, G. and Gollwitzer, P.M. (2010) Intervention effects of information and self-regulation on eating fruits and vegetables over two years . *Health Psychology* , 29, 274–283. <http://dx.doi.org/10.1037/a0018644>

Taylor, S.E. and Armor, D.A. (1996) Positive illusions and coping with adversity . *Journal of Personality* , 64, 873–898. <http://dx.doi.org/10.1111/j.1467-6494.1996.tb00947.x>

Taylor, S.E. and Brown, J.D. (1988) Illusion and well-being: A social psychological perspective on mental health . *Psychological Bulletin* , 103, 193–210. <http://dx.doi.org/10.1037/0033-2909.103.2.193>

Taylor, S.E. and Gollwitzer, P.M. (1995) The effects of mindsets on positive illusions . *Journal of Personality and Social Psychology* , 69, 213–226. <http://dx.doi.org/10.1037/0022-3514.69.2.213>

Taylor, S.E., Kemeny, M.E., Reed, G.M., Bower, J.E. and Gruenewald, T.L. (2000) Psychological resources, positive illusions, and health . *American Psychologist* , 55, 99–109. <http://dx.doi.org/10.1037/0003-066X.55.1.99>

Watt, H.J. (1905) Experimentelle Beiträge zu einer Theorie des Denkens (Experimental analyses of a theory of reasoning) . *Archiv für die gesamte Psychologie* , 4, 289–436.

Weinstein, N.D. (1984) Why it won't happen to me: Perception of risk factors and susceptibility . *Health Psychology* , 3, 431–457. <http://dx.doi.org/10.1037/0278-6133.3.5.431>

Wicklund, R.A. and Gollwitzer, P.M. (1982) Symbolic Self-completion . Hillsdale, NJ: Erlbaum.

Wong, E.M., Galinsky, A.D. and Kray, L.J. (2009) The counterfactual mindset: A decade of research . In K.D. Markman, ed. , W.M.P. Klein, ed. , and J.A. Suhr (eds), *Handbook of Imagination and Mental Simulation* , pp. 161–174. New York: Psychology Press.

[p. 546 ↓]

<http://dx.doi.org/10.4135/9781446249215.n26>

Handbook of Theories of Social Psychology

Self-Control Theory

Contributors: Walter Mischel

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Self-Control Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n26>

Print pages: 1-23

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n26>

[p. 1 ↓]

Chapter 26: Self-Control Theory

WalterMischel

Abstract

The self-control theory my colleagues and I developed evolved slowly over many decades, intended to integrate work on two closely interconnected questions of both my scientific and personal life. From the start of my career I was driven by these two questions, and I felt that the answer to each hinged on the answer to the other. First, given the power of the situation, demonstrated so often by social psychologists, how do individuals manage, at least sometimes, to inhibit and control their impulsive automatic responses to powerful situational pressures, overcoming “stimulus control” with “self-control”? Second, given the great variability one sees in what anyone does and thinks and feels across different situations, what are the consistencies that distinctively characterize individuals over the life course? In this chapter I discuss some of the empirical labors and surprising discoveries, as well as ideas, and the good luck, that ultimately allowed some answers to both these questions that seem – for the moment – to reasonably fit the integrative and still-evolving theory that emerged.

First Steps

It didn't turn out the way I expected. Half a century ago, I flew on a prop plane out of a cold, slushy, Columbus Ohio January, and landed in sun-drenched, beautiful blue-sky Trinidad, eager to observe and study spirit possession as practiced in the Orisha religion, a blend of African and Catholic beliefs, by a group then known as Shango. The chance to leave Columbus for travel to what then felt like exotic rum and Coca-Cola places on palm-lined beaches outside the tourist routes, still under British colonial rule, was irresistible. It also allowed a break from my graduate training, working towards a

clinical psychology PhD at Ohio State from 1953 to 1956. During those years I went repeatedly to Trinidad with my then wife, Frances (now Frances Henry), who was, in the early 1950s, a doctoral student in anthropology. We hoped to find connections between what people do in their daily lives, in which they served the most menial roles at the bottom of their stratified, still-British-colonial society and what they did and became when “possessed” (Mischel and Mischel, 1958). My clinical experiences at Ohio State were already making me worried about the value of projective measures for making decisions in the mental hospital setting. But I was still hopeful about their potential for exploring what goes on at the fantasy level, and thought it was worth a try. Armed with my Rorschach inkblot cards (and a sketch pad), we headed for Trinidad for several summers from 1955 to 1958.

[p. 2 ↓]

From Trinidad and Spirit Possession to Delay of Gratification

A Brief Anthropology Venture

We found Shango practiced in a small village in the southern tip of Trinidad, and befriended its leader, Pa Neezer, who congenially welcomed us to live in one of his little houses as observers to study and understand their ceremonies and practices. The participants in Shango were cooperative and eager to please. This became clear in their responses to my projective tests. Their answers often had more to do with what they thought might interest me (e.g., plots from current American films playing in the next town) than with their inner lives. If the stories they spun had some connections to their inner states, it was beyond me to discern them. They certainly seemed unconnected to what they did in front of my eyes while they were possessed by the spirits that “rode” them during the Shango ceremonies (described in Mischel and Mischel, 1958). I soon put the tests away and started looking at what might be going on around us.

In the Shango ceremonies, stretching over several days and nights, laborers and domestic servants of the British by day became at night possessed by spirits that were a mix of Catholic saints and African gods, and danced in hypnotized trance-like states, with the irresistible drums pounding, and the rum bottles passing. As I found myself struggling to resist hurling myself into the dance I realized that the participant–observer balance was tipping fast, and that while keeping my eyes glued to the scene, I needed to turn my work in other directions.

Small Candy Now versus Big Candy Later

My transition from resisting the impulse to jump into the dance and instead to invent what decades later the media call the “marshmallow test” took more than a dozen years. It began when I started talking to our neighbors in the village and listened, really listened, as they talked about their lives. The inhabitants in this area of the island were of either African or East Indian background, each group living in its own enclave, on different sides of the same long street. It did not take much listening to hear a recurrent theme in how they characterized each other. The Africans, the East Indians said, were just pleasure-bent, impulsive, eager to have a good time, and live in the moment while never planning or thinking ahead about the future. Reciprocally, the Africans saw their East Indian neighbors as just working for the future and stuffing their money under the mattress without ever enjoying today.

I started to examine these observations in the local schools with young children from both ethnic groups. In their classrooms, I administered a variety of measures that ranged from such demographic descriptors as father's presence–absence in the home, to trust expectations, achievement motivation, diverse indices of social responsibility, and intelligence. At the end of each of these sessions I gave them choices between little treats (a tiny chocolate bar, a small notepad) that they could have immediately or a much bigger, better one that they would get the following week.

Consistent with the stereotypes the two groups had about each other, the black Trinidadian kids generally preferred the immediate rewards and those from East Indian families chose the delayed ones much more often (Mischel, 1961a, 1961b, 1961c). I wondered if the kids who came from homes with absent fathers – then common in

the black families in Trinidad, very rare for the other group – might have had fewer experiences with male social agents who kept their promises. If so, those children would have less trust; that is, a lower “expectancy” that the male, and on top of it white, visitor would show up with the promised delayed reward. I had learned about the importance of expectancies from my Ohio State mentor, Julian [p. 3 ↓] Rotter, and was impressed by his social learning theory, and the expectancy construct also was the topic of my doctoral dissertation. So I controlled for the effect of father absence, and was delighted to see that the differences between the ethnic groups disappeared. These findings pointed to the important role that outcome expectancies and beliefs play in goal commitment. People are likely to attempt to exercise self-control (and forego the “bird in the hand”) only if they trusted that the delayed larger one (“in the bush”) would materialize (Mischel, 1974).

This was the beginning of the studies that identified some of the main determinants of such choice behavior and what later became known as “temporal discounting” – a major topic in current behavioral economics (e.g., Mischel, 1961a, 1961b, 1961c; Mischel and Gilligan, 1964; Mischel and Metzner, 1962; Mischel and Mischel, 1958). These studies showed significant correlations between dominant choice preferences for the immediate rewards and, for example, juvenile delinquency in adolescence, lower social responsibility ratings, less resistance to temptation in experimental situations, lower achievement motivation, and lower intelligence (see Mischel, 1974, for a summary).

Expectancies (Trust) and Values

Goal commitment does not just depend only on peoples’ trust expectations. It is also influenced by the subjective value of the rewards in the situation. Through temporal discounting mechanisms, rewards that are delayed have less value than equivalent rewards that are immediately available (Ainslie, 2001; Loewenstein et al., 2003; Rachlin, 2000). Therefore we expected, and found, that the longer the future rewards were delayed, the less likely it was that children would choose to wait for them (Mischel and Metzner, 1962). Thus, goal commitment in delay of gratification is enhanced with the relative magnitude of the delayed reward and decreases as the required time it takes to attain the reward increases (Mischel, 1966, 1974).

At that point, and consistent with utility theories in economics as well as in psychology, the findings indicated that the choice to wait for a larger but delayed reward is determined largely by an expectancy-value mechanism (discussed in Mischel, 1974; Mischel and Ayduk, 2004). In short, a person must value the delayed reward enough to commit to pursuing it, must believe that they possess the ability to successfully exert self-control should they choose to do so (e.g., Bandura, 1986; Mischel and Staub, 1965), and trust that they will receive the valued delayed reward upon successfully fulfilling their goal. And the delayed rewards must matter to them enough to overcome the temporal discounting effects. Further, we found that these choices could be modified, in either direction, by exposure to high prestige peer models who modeled choice preferences opposite to those of the subject. These changes were still evident several months later when the participants were again tested in a new situation (Bandura and Mischel, 1965).

A Window for Watching Willpower

Willpower: The Scorned Fiction of Behaviorism

When I initiated my self-control research, behaviorism was the dominant theory, or anti-theory, regnant within American psychological science in the 1950s and early 1960s. Academic psychology was still deep into “positivism” (not to be confused with the current “positive psychology”), well before the cognitive revolution, and dominated by radical behaviorism and Skinner’s focus on “stimulus control” and the power of reinforcement. Such concepts as “self-control” and “self” were dismissed as naïve unscientific fictions, and even the word “willpower” was unspeakable in academic circles.

[p. 4 ↓]

The ridiculed willpower “fiction” became my focus and research agenda as I watched my three closely spaced daughters each morph in the first few years of life, from mostly

giggling and gurgling or screaming and sleeping, to becoming people with whom one could have fascinating, thoughtful conversations. Most amazing to me, sometimes they could even sit still for a while to wait for things they wanted that took some time or effort to get. As I tried to make some sense out of what was unfolding in front of me at the kitchen table I mused that in behaviorism as well as in economics the explanatory keys for most human behavior, including what was happening to my children, were rewards. But I did not have a clue about how rewards enable voluntary delay of gratification and "willpower," a term that as a psychologist I even now put into quotes. And trying to understand how that happens became a lifelong obsession.

The decision before entering the restaurant to forego dessert, and the ability to stick to it when the pastry temptations are flashed in front of one's eyes, often are unconnected. The firmest New Year's resolutions easily break before January ends, and the tobacco addict who dumps his cigarettes into the garbage in self-disgust, vowing to quit forever, may be frantically searching for them three hours later. Therefore the question I kept asking myself became: After the choice to delay has been made, the good intention formed and declared at least to oneself, what allows it to be realized? And how does this ability develop in the young child?

The Marshmallow Test

To go from speculating to empiricism we needed a method to study delay of gratification ability when the young child begins to have it around preschool age. Happily, the newly established Bing Nursery School at Stanford University, with its big one-way glass observation windows, was the ideal laboratory, and as a newly arrived (in 1962) faculty member I was thrilled to use it. In the next dozen years, my students, notably including Ebbe Ebbesen, Bert Moore, and Antonette Zeiss (but many others also played important parts) and I came up with the preschool "delay of immediate gratification for the sake of delayed but more valued rewards paradigm," in the media later called more simply, albeit incorrectly, *the marshmallow test*.

Typically a preschooler is shown some desired treats; for example, small marshmallows or (more often) little pretzel sticks, or cookies, or tiny plastic toys. The child faces a conflict: wait until the experimenter returns and get two of the desired treats, or ring

a bell and the experimenter will come back immediately – but then the child gets only one treat. After the child chooses to wait for the larger outcome, he or she is left alone, waiting while facing both treats, and the measure is the seconds of delay before settling for the one or waiting the full time to get the two (e.g., after 15 minutes). The delay soon becomes difficult and frustration grows quickly. As waiting for the chosen goal drags on, the child becomes increasingly tempted to ring the bell and take the immediately available treat.

This situation has become a prototype for studying the conflict between an immediate smaller temptation and a higher-order but delayed larger goal, the bigger treat that will come later (when the experimenter returns). In this type of situation, my students and I studied hundreds of preschoolers in the Stanford University community, both with experiments and through direct observation, with follow-up studies that are still in progress (e.g., Mischel et al., *in press*) We began with a series of experiments designed to see how the mental representation of the rewards in the choice situation influence the ability to resist impulsive responding and to continue to wait or work for the chosen delayed but more valuable outcomes, as described in later sections.

The experiments were designed to identify the mental processes that allowed some [p. 5 ↓] people to delay gratification while others simply couldn't. I had no reason to expect that seconds of waiting time for a couple of marshmallows or cookies at age four years would predict anything worth knowing about years later. In fact, there was every reason to not expect that since successful attempts to predict long-term consequential life outcomes from psychological tests very early in life were proving to be exceptionally rare (Mischel, 1968). But occasionally I did ask my three daughters, who all had attended the Bing school, how their friends from nursery school were doing as the years passed. Far from systematic follow-up, this was just idle dinnertime conversation, as I asked them: "How's Debbie? How's Sam doing?" By the time the kids were early teenagers I noticed what looked like a possible link between the preschoolers' scores on the "marshmallow test" and the informal judgments about their academic and social progress when I asked my daughter informally to rate their friends on a scale of zero to five. Comparing these ratings with the original data set, a clear correlation was emerging, and I realized I had to do this seriously.

An Unexpected Dividend: Preschool Delay Predicts Long-Term Outcomes

Beginning in 1981, my students and I sent out a questionnaire to the reachable parents, teachers, and academic advisers of preschoolers who had participated in the delay research and who by then were in high school. We asked about all sorts of behaviors and characteristics that might be relevant to impulse control, ranging from their ability to plan and think ahead to their skills and effectiveness in coping with personal and social problems (e.g., how well they got along with their peers). We also requested and obtained their Scholastic Aptitude Test (SAT) scores from the Educational Testing Service. It soon became evident that there were long-term differences between the preschoolers who were high and low delayers, and we therefore continued to examine them systematically as they developed over many years, and still do so as they reach their mid-forties.

Stumbling into the Bing Longitudinal Study

Participants in the Bing Longitudinal Study come from a sample of more than 300 participants who were enrolled in Stanford University's Bing preschool between 1968 and 1974. Since then, we assessed the ability of these participants to pursue long-term goals in the face of immediate temptation once every decade since the original testing. They now have reached their late thirties and early to mid forties, and information about their life outcomes, such as their occupational, marital, physical health, and mental health status are continuing to become available. The findings have surprised us from the start, and they continue to do so. For example, preschoolers who delayed longer relative to other participants earned much higher SAT scores (on average about 200 points higher) and exhibited better social–cognitive and emotional coping in adolescence (Mischel et al., 1988, 1989; Shoda et al., 1990).

When the high delayers became adults, most continued to have better cognitive–social functioning and better educational and economic life outcomes than their low-delaying peers (e.g., Ayduk et al., 2000; Mischel and Ayduk, 2004; Mischel et al., 1988; Shoda

et al., 1990). The high delay individuals also were buffered against the development of diverse mental health problems: they used cocaine/crack less frequently, were less likely to suffer from low self-esteem and self-worth (Ayduk et al., 2000), and had fewer features of borderline personality disorder than matched controls with similar dispositional vulnerability (Ayduk et al., 2008).

To be sure that what we were discovering in these long-term correlates was not [p. 6 ↓] restricted to the Bing cohort, we also conducted longitudinal studies with similar measures in a variety of other cohorts and demographic populations. We obtained closely parallel findings with children from the toddler center at Barnard College in New York (e.g., Eigsti et al., 2006; Sethi et al., 2000), and with middle-school children in the South Bronx, New York (e.g., Ayduk et al., 2000; Mischel and Ayduk, 2004), as well as with children and adolescents in a summer residential treatment program for youths at high risk for problems of aggression/externalization and depression/withdrawal (e.g., Mischel and Shoda, 1995; Rodriguez et al., 1989). For example, spontaneous use of self-control strategies in the delay task (e.g., looking away from the rewards, in this case M&M candies, and using self-distraction) predicted reduced verbal and physical aggression as directly observed over six weeks in the summer camp study (e.g., Rodriguez et al., 1989; Wright and Mischel, 1987, 1988).

Decomposing the Ability to Delay Gratification/Impulse Control

The fact that the marshmallow test's long-term predictive power turned out to be substantial made me even more eager to understand the cognitive-affective mechanisms that underlie the individual differences in self-control tapped by the test. We did those experiments at Bing in the 1970s, before the longitudinal research was launched, with the hope of finding the mental mechanisms that enable delay of gratification.

Conceptual Roots

Initially, I was guided by the idea that delay becomes easier when the desired gratification can be visualized (Mischel et al., 1972). That hypothesis was based on Freud's (1911/1959) classic idea that delay of gratification becomes possible when the young child creates a mental ("hallucinatory" was Freud's phrase) image of the object of desire (e.g., the mother's breast). In Freud's view, the mental representation of the object allows mental "time binding" and enables the transition from primary process thinking to delay and impulse inhibition (Rapaport, 1967). Using very different language a similar idea came from experiments by researchers working at the behavioral-conditioning level. Their research (e.g., Berlyne, 1960; Estes, 1972) suggested that when animals learn, their approach behavior toward a goal is maintained by "fractional anticipatory goal responses" that cognitively represent the desired rewards. These anticipatory representations sustain the rat's goal pursuit, for example, as it tries to find its way back to the food at the end of a maze in a learning task (Hull, 1931). Again, the prediction was that focusing attention on the delayed rewards should reinforce one's ability to sustain delay gratification in order to fulfill goal pursuit. In the first experiments on delay with four-year-olds we examined these ideas, predicting that waiting would be longer if the rewards were made available for attention during the delay period. The results turned out to be the direct opposite of what we expected.

We got these upsetting results from a series of experiments to explore the role that attention to the rewards plays in self-control (Mischel and Ebbesen, 1970). With that goal, we varied whether or not reward items were available for attention while children were waiting in the delay of gratification paradigm. In one condition, children waited with both the immediately available and the delayed reward exposed in full view. In a second condition both options faced the child, but were concealed from attention by an opaque cover positioned over them. In two other conditions either the delayed reward alone or the immediately available reward alone was exposed during the delay period. On average, children waited more than [p. 7 ↓] 11 minutes when none of the rewards were exposed, but waited only a few minutes when any of the rewards – either both rewards, just the delayed reward, or just the immediately available reward – were available to attention. Directly contradicting the predictions coming from both from

the psychodynamic and animal learning traditions, the results showed that focusing attention on a desired stimulus decreased the ability to delay gratification.

To try to figure out what might be going on in the heads of the preschoolers as they tried to wait in the marshmallow test, chatting with my daughters and doing some playful but serious mini-experiments with them gave me many hypotheses. And for many hours my students and I simply observed preschool children at the Bing school in their “game room” through the windows of the one-way glass while they were struggling to wait to get the more valuable treat later or rang the bell to get the less valuable one immediately. We saw that the kids who managed to delay were doing anything they could to distract themselves from the rewards and reduce their frustration while continuing to wait; for example, by fidgeting, squirming, hiding their eyes to not see the temptations, kicking the table, playing with their toes and fingers, picking their noses and ears in elaborately imaginative ways, singing little songs they invented (“Oh this is my home in Redwood City”), and so on.

If shifting attention away from the rewards to reduce the temptation is what matters, then distracting children from focusing on the rewards should have the same effect as removing the rewards from view. That's just what was found. In one experiment, for example, we provided children with a distracting toy (a Slinky) to play with while they tried to wait, facing the rewards exposed on the table in front of them (Mischel et al., 1972). In this condition more than half of the children waited the full amount of time until the experimenter returned indicating that the experiment was over (15 minutes). In contrast, none of the children who were left waiting for the exposed rewards without the distracter toy were able to do so. In another experiment, the same effect of distraction on delay times was found when children were cued to think about fun thoughts while they waited: “While you're waiting, if you want to, you can think of mommy pushing you on a swing at a birthday party.” Similar to the Slinky condition, more than half of the children who were cued to distract themselves with fun thoughts waited until the experimenter returned and indicated that the experiment was over (Mischel et al., 1972). Of course not all distracters were equally effective. Unsurprisingly, when the distracting object was not appealing, for example, instructing individuals to think about sad thoughts, then attention was diverted back to the stimulus and delay of gratification was undermined. To effectively keep attention away from the temptations in the situation, attention to the distracter must itself be reinforcing.

Developmental and social–cognitive research points at similar attentional processes in regulating negative affect and behavior. For example, eye-gaze aversion, flexible attention shifting, attention focusing, and resistance to attentional interference are related to reduced impulsivity and anger even in early childhood (Eisenberg et al., 2002; Johnson et al., 1991; Posner et al., 1997). Likewise, social–cognitive research indicates that whereas processes such as emotion-focused rumination maintain and prolong negative affect, self-distraction may be an effective strategy to assuage negative mood (Nolen-Hoeksema, 1991; Rusting and Nolen-Hoeksma, 1998).

Reappraisal Processes: From Hot to Cool

Strategically focusing attention away from a desired stimulus is an effective way of facilitating adaptive self-control in the face of temptation, but that option often is not available or not sustainable. Consider, for example, the dieting pastry chef who has [p. 8 ↓] sworn off eating chocolate, yet has to make delectable chocolate cakes for dessert each night, creating one potential conflict after another.

In the late 1960s, we began to test systematically how alternative ways of mentally representing the stimulus influence the emotions and behaviors of the children during their self-control efforts. We drew on a distinction that had been made in the research literature between two different aspects or features of a stimulus: its “hot” motivational, consummatory, arousing, action-oriented, or motivating “go” features; and its informational, “cool,” cognitive cue or discriminative stimulus “know” functions (Berlyne, 1960; Estes, 1972). Given this distinction, Mischel and Moore (1973) reasoned that when a child thinks about the rewards in front of them as “real,” attention is placed on their hot, arousing, consummatory features, which should in turn elicit the motivating effects of the stimulus, making delay of gratification more difficult, and leading quickly to the “go” response: ring the bell, get the treat now. In contrast, we predicted that thinking about the rewards in terms of their cooler, more abstract features should allow the child to focus on the reward without activating consummatory trigger reactions. For example, mentally representing the rewards as pictures emphasizes their cognitive, informational features rather than their consummatory features. Therefore we speculated that this kind of “cool” mental transformation would reduce the conflict between wanting to wait and wanting to ring the bell by shifting attention away from the

arousing features of the stimulus and onto their informative meaning (also see Trope and Liberman, 2003).

To test this prediction, Bert Moore and I presented one group of children in the delay of gratification task with slide-presented lifesize pictures of the rewards, formally called "iconic representations." The hypothesis again was that the pictures of the rewards would be relatively more abstract than the actual rewards, and thus the temptation to reach for them should be attenuated. These iconic representations were pitted against the presence of the real rewards themselves during the delay period. As predicted, exposure to the pictures of the images of the rewards significantly increased children's waiting time whereas exposure to the actual rewards decreased delay time (Mischel and Moore, 1973).

In one study, children were faced with actual rewards while they tried to wait, but this time the experimenter cued them in advance to "just pretend" that they were pictures: "Just put a frame around them in your head"(Moore et al., 1976). In a second condition, the children were shown pictures of the rewards but this time asked to think about them as if they were real. The children were able to delay almost 18 minutes when they pretended that the rewards facing them were pictures. In contrast, they were able to wait less than 6 minutes if they pretended that pictures in front of them were real rewards. As one child put it when asked in the postexperimental inquiry how she was able to wait so long: "You can't eat a picture."

In a study with Nancy Baker we identified the types of cognitive reconstrual that facilitate the ability to delay gratification. In this study, we cued children to represent the rewards available in front of them in terms of either their cool informational or hot consummatory features. For example, children in the cool focus condition who were waiting for marshmallows were cued (or "primed" in current terminology) to think of them as "white, puffy clouds." Those waiting for tiny stick pretzels were cued to think of them as "little brown logs." In the hot ideation condition, the instructions cued children to think about the marshmallows as "yummy and chewy" and the pretzels as "salty and crunchy." As expected, when children thought about the rewards in hot terms, they were able to wait only 5 minutes, whereas when they thought about them in cool terms, delay time increased to 13 minutes (Mischel and Baker, 1975).

[p. 9 ↓]

Summary: Hot versus Cool Focus

In sum, attention to the rewards may either make delay easier or harder, depending on whether the focus on the consummatory (hot, emotional) or nonconsummatory (cool, informational) features of the temptations. A nonconsummatory focus on the rewards can help self-imposed delay even more than comparable distractions; a consummatory focus makes delay exceedingly difficult. How the rewards are represented cognitively in this regard, crucially influences the duration of delay in opposite ways (Mischel et al., 1989). This now seems evident, but 35 years earlier, when behaviorism prevailed and the cognitive revolution was in its infancy, it was startling. For me it was the tipping point, from a focus on external stimulus control to internal self-control and the conditions that enable it as the person interacts with the social world.

The experiments in the late 1960s and early 1970s gradually, step by step, made it clear that the crucial determinant of the young child's ability to delay immediate gratification was not the rewards faced in the situation, as earlier theories had suggested. Instead, and contradicting the expectations both of classic behaviorism and of Freud, what mattered was exactly how they were represented mentally (Mischel, 1974; Mischel et al., 1989). The duration of delay depended on specific types of "hot" or "cool" mental representations, and the precise ways in which attention was deployed during the delay interval (e.g., Mischel and Baker, 1975; Mischel and Moore, 1973; Mischel et al., 1972; Peake et al., 2002). The best news for me was that children could be primed to change the representation from hot to cool, making it much easier for them to exert self-control when needed. If they could be primed by the experimenter in the lab, perhaps they also could learn to activate the needed strategies themselves and to plan to use them in pursuit of their own goals in everyday life.

Making Self-Control Automatic: Plans

To exercise self-control effectively when it's needed in vivo in "hot" situations, needed strategies to cool the temptations and maintain adaptive delay behavior have to be

activated virtually reflexively. That requires shifting from effortful or “volitional” control to automatic and virtually reflexive activation when they are needed. To explore the mechanisms enabling such a shift, Charlotte Patterson and I examined how different types of plans and rehearsal strategies facilitate preschool children’s ability to resist temptation (Mischel and Patterson, 1976; Patterson and Mischel, 1976).

In these experiments preschool children were motivated to work on a long, repetitive task (sticking pegs into holes) in order to get attractive rewards later, and were warned that a “Mister Clown Box” might tempt them to stop working on the task. A temptation-inhibiting plan suggested that they direct attention away from Mister Clown Box; a task-facilitating plan suggested that they direct their attention toward continuing to work on the task. Some children received both plans, another group were given no plans, a third received only the temptation inhibiting plan, and a fourth only the task-facilitating plan. After the self-instructional manipulations, the child was left alone to work while the Clown Box performed a standard routine designed to tempt the child to stop working (e.g., “Please, please come talk to me … I have big ears and love to have children talk into them and tell me what they think and want.” Mister Clown Box, who had a colorfully painted clown face and display windows, exhibited tempting toys placed on a rotating drum in two windows, as he urged the child to “come talk with me and play with my toys.” Dependent measures assessed the amount and rate of “work” completed, and the allocation of attention while dealing with the temptations.

[p. 10 ↓]

The children struggled, often desperately, to resist the temptations, pleading with Mister Clown Box (e.g., “Don’t talk to me,” “Stop that,” “Please don’t bother me”). Their spontaneous effective strategies were very similar in intent to the temptation-inhibiting plan suggested by the experimenter, in that both seem designed to suppress the distracting stimuli in the child’s environment. The results of this and related studies made clear that the effective plan was the one with specific self-instructions to resist the temptation when it occurred, and that having such plans available and accessible greatly facilitated persistence in goal-directed activity (Mischel and Patterson, 1976; Patterson and Mischel, 1976). The importance of such plans is now fully recognized by Peter Gollwitzer and colleagues in their persuasive and systematic studies showing the value of specific “if–then” implementation plans in the actualization of effective

self-control strategies under stressful conditions, and further clarifying the relevant mechanisms (Gollwitzer, 1999).

Hot/Cool System Interactions in Self-Control

The important long-term correlates of the marshmallow test, and the clear findings from the experiments that helped identify the ability it tapped, led me to become interested in a more formal way to conceptualize those results. That called for a model of impulse control that could be integrated within the broader Cognitive–Affective Processing System (CAPS) that Yuichi Shoda and I had designed for understanding the expressions of stable individual differences in person-situation interactions (Mischel, 1973; Mischel and Shoda, 1995). With that goal, Janet Metcalfe and I proposed “a two-system framework for understanding the processes that enable—and undermine—self-control or ‘willpower’ in the execution of one’s intentions, as exemplified in the delay of gratification paradigm” (Metcalfe and Mischel, 1999). We postulated two closely linked systems: a cool cognitive ‘know’ system, and a hot emotional ‘go’ system (see also Metcalfe and Jacobs, 1996, 1998).

The hot system is an automatic system that responds reflexively to trigger features in the environment, both positive and negative, and elicits automatic, aversive, fight-and-flight reactions as well as appetitive and sexual approach reactions. It consists of relatively few representations which, when activated by trigger stimuli, elicit virtually reflexive avoidance and approach reactions. The cool system, on the other hand, is conceptualized as a controlled system that is attuned to the informational, cognitive, and spatial aspects of stimuli. It consists of a network of informational, cool nodes that are elaborately interconnected to each other, and generate rational, reflective, and strategic behavior. Whereas the hot system is conceptualized as the basis of emotionality, the cool system is thought to be the basis of self-regulation and self-control.

This hot/cool system idea is of course at least metaphorically related in its historical roots to Freud’s conception of the id as characterized by irrational, impulsive urges for immediate wish-fulfillment, and its battles with the rational, logical, executive ego. The

difference is that what has been learned from research on this topic over the course of the past century, not least the break-through in methods for imaging activity in the brain, we now can specify more clearly the cognitive and emotional processes, and even the neural process, that underlie these two systems and their interactions to enable effective self-regulation (e.g., Mischel et al., *in press*).

The regions of neural activity underlying these different systems currently remain a vigorously pursued topic of research (for review see Kross and Ochsner, 2010; also see Lieberman, 2007; Mischel et al., *in press*; Ochsner and Gross, 2005). Collectively the [p. 11 ↓] findings point to the amygdala – a small, almond-shaped region in the forebrain thought to enable fight or flight responses – as critically involved in hot system processing (Gray, 1982, 1987; LeDoux, 2000; Metcalfe and Jacobs, 1996, 1998). This brain structure reacts almost instantly to stimuli that individuals perceive as arousing (Adolphs et al., 1999; LeDoux, 1996, 2000; Phelps et al., 2001; Winston et al., 2002), immediately cueing behavioral, physiological (autonomic), and endocrine responses. The cool system, in contrast, seems to be associated with prefrontal and cingulate systems involved in cognitive control and executive function (e.g., Jackson et al., 2003; Ochsner and Gross, 2005).

The two systems continuously interact with each other, and with the stimuli in the particular context, producing the individual's subjective experiences and behavioral responses (also see Epstein, 1994; Lieberman et al., 2002). Hot representations and cool representations that have the same external referent are directly connected to each other, and link the two systems (Metcalfe and Mischel, 1999; see also Metcalfe and Jacobs, 1996, 1998). Thus hot representations can be evoked by the activation of corresponding cool representations. For example, an abusive man can become enraged by conjuring up a fantasy in which he finds his partner cheating on him. Likewise, hot representations can be cooled through the activation of cool system cognitive processes (e.g., attention switching, reconstrual). Thus the same abusive man can calm himself down by distracting himself or by recognizing that his fantasy is his own self-created fiction. Self-control becomes possible to the extent that cooling strategies are generated by the cognitive cool system to reduce hot system activation. While the particulars are different, the basic mechanisms are no different than those that regulate the child's ability to self-control in the marshmallow test. For the man in the

example, the delayed important consequences are preservation of the relationship; for the child in the preschool delay situation, attainment of the two marshmallows.

Effects of Stress

The balance of hot/cool system processing is influenced by several factors. The most important determinant of hot/cool system balance in adults tends to be stress. When stress levels are high, the cool system becomes deactivated and the hot system dominates. This makes complex thinking, planning, and remembering virtually impossible, ironically just when it may be most needed. When stress levels jump from low to very high, as in life-threatening emergency conditions, responding tends to be reflexive and automatic. That was probably highly adaptive in earlier evolutionary times: when an animal's life is threatened in the jungle, quick responses driven by innately determined stimuli may be essential. But when humans quarrel angrily at the breakfast table, such automatic reactions undermine rational efforts at constructive self-control.

Developmental Level

Age and maturation matter. Early in development young children are primarily under stimulus control, because they have not yet developed the cool system structures needed to regulate hot system processing. The hot system develops and dominates early in life, whereas the cool system develops later (by age four) and it becomes increasingly dominant over the developmental course. These developmental differences are consistent with the differential rates of development of the relevant brain areas for these two systems (for reviews see Eisenberger et al., 2004; Rothbart et al., 2004). As the cool system develops, children become increasingly able to generate cooling strategies to regulate impulses (Mischel et al., 1989). These developmental changes also may underlie the [p. 12 ↓] greater vulnerability to the effects of stress and traumas early in life.

Towards an Integrative Caps Self-Control Theory

The important and stable long-term differences in self-control tapped by the marshmallow test, and the cognitive–affective mechanisms that underlie those individual differences, coexist with the fact that self-control behavior, like all social behavior, is expressed in highly contextualized if–then situation-specific ways. And that has crucial implications for the integrative CAPS theory of self-control that my colleagues and I developed (e.g., Mischel and Shoda, 1995).

The If–Then Contextualized Expressions of Self-Control

It is true that, on average, the high delay group in the Bing lifespan study looks very different from the low delay group in follow-ups conducted at roughly ten-year intervals. But when we examine these differences closely, we also see that within each group, and within each individual, there is equally impressive variability. Well-known examples of such variability abound in daily life. Former president Bill Clinton was clearly high in his average overall ability to exert self-control and delay gratification. Without it, he could never have become president of the US, not to mention a Rhodes scholar and a Yale-trained lawyer. Yet, evidence for his systematic failures to exert self-control came in the painful details of his descent towards impeachment (Morrow, 1998).

Less publicized, but even more surprising for many, was the fall of Sol Wachtler, Chief Judge of the State of New York and the Court of Appeals, to incarceration as a felon in federal prison. Judge Wachtler was well known for advocating laws to make marital rape a punishable crime, and he was deeply respected for his landmark decisions on free speech, civil rights, and the right to die. After his mistress left him for another man, however, Judge Wachtler spent 13 months writing obscene letters, making lewd phone calls, and threatening to kidnap her daughter. His descent from the court's bench as the model of jurisprudence and moral wisdom to federal prison testifies that smart people

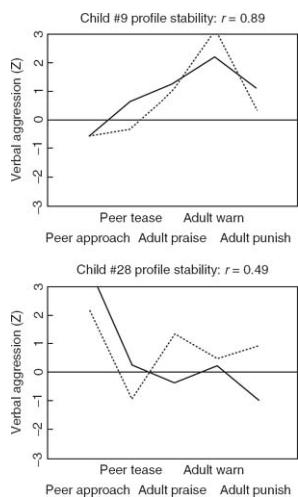
are not necessarily consistently so across different areas of their lives (e.g., Ayduk and Mischel, 2002). As observers of human behavior have long known, even “on average” adaptively controlled lives are not without their surprising failures to exert such control at crucial times.

These everyday observations are supported by extensive research that examined closely the consistency of social behavior as it actually unfolds across diverse situations (e.g., Mischel, 1968, 2004, 2009; Mischel and Peake, 1982a; Mischel and Shoda, 1995). To illustrate, in the Carleton College field study, behavior relevant to conscientiousness in college was observed *in vivo* over multiple situations and occasions (Mischel and Peake, 1982a, 1982b). Each of the 63 participating college students was observed repeatedly in various situations on campus relevant to their conscientiousness in the college setting. The undergraduates themselves supplied the contexts or situations they considered relevant. Based on this information from pretests, the students’ conscientiousness was sampled in diverse situations, such as in the classroom, in the dormitory, or in the library, and these assessments were repeated over multiple occasions in the course of the semester. The directly observed actual consistency correlation in their cross-situational behavior was on average between 0.08 (for single behaviors) to 0.13 for reliable aggregates of the single behaviors within each of the 19 types of conscientiousness sampled. Thus while the correlations were not zero-order, they made clear that an individual may be highly conscientious in one type of situation, and much less conscientious than [p. 13 ↓] most people in another type of situations, even if both types seem highly similar (Mischel and Peake, 1982a).

Consistent with these findings, and flying in the face of the core assumptions of traditional personality psychology that personality traits are expressed consistently across diverse situations, my monograph, *Personality and Assessment* (1968), called attention to the highly contextualized, situation-specific expressions of individual differences in social behavior. The conscientious man at work may be a scoundrel in his private life; the aggressive child at home may be less aggressive than most when in school; the man exceptionally hostile when rejected in love may be unusually tolerant about criticism of his work; the one who shakes with anxiety in the doctor’s office may be a calm mountain climber; the business entrepreneur may take few social risks. And 40 years later a great deal of behavioral evidence continues to support this perspective (e.g., Mischel, 2009; Orom and Cervone, 2009; Van Mechelen, 2009).

But while individual differences are rarely expressed in consistent cross-situational behavior across widely different situations, the new discovery is that consistency is found in distinctive but stable patterns of if–then situation-behavior relations. These patterns of variability form contextualized, psychologically meaningful “personality signatures” (e.g., “she does A when X, but B when Y” that are stable over time). Such behavioral signatures were first revealed in a massive fine-grained observational study of social behavior across multiple repeated situations over time in a summer camp for children and adolescents (Mischel and Shoda, 1995). We found that individuals who were similar in average levels of behavior, for example in their aggression, nevertheless differed predictably and dramatically in the types of situations in which they aggressed; that is in their if–then situation-behavior signatures.

Figure 26.1 Individual if–then situation-behavior signatures for two children. Their aggressive behavior was observed in five different situations many times. Half of the observations are shown as dotted lines; half as solid lines. Profile stability is the correlation between the two sets of observations. (From Shoda, Y., Mischel, W., and Wright, J.C. (1994) Intra-individual stability in the organization and patterning of behavior: Incorporating psychological situations into the idiographic analysis of personality. Journal of Personality and Social Psychology, 67, 674–687, fig. 1. © 1994 by the American Psychological Association. Adapted with permission.)



As [Figure 26.1](#) shows, each child showed a distinctive and stable if–then situation-behavior pattern or “profile” that distinguished him or her characteristically (Shoda et al., 1994). Even when two children are equal in their overall aggressive behavior, for example, the one who regularly becomes aggressive when peers try to play [[p. 14 ↓](#)] with him is quite different from the one who expresses aggression mostly to adults who try to control him. In short, stable situation-behavior personality signatures, not just stable levels of average overall behavior, characterize individuals, and this is the case for the expressions of conscientiousness and self-control as much as for other individual differences.

These stable signatures of personality have now been extensively documented in various studies of observed behavior as it unfolds over time and across diverse situations (e.g., Andersen and Chen, 2002; Borkenau et al., 2006; Cervone and Shoda, 1999; Fournier et al., 2008; Morf and Rhodewalt, 2001; Moskowitz et al., 1994; Shoda and LeeTiernan 2002; Vansteelandt and Van Mechelen, 1998; Van Mechelen, 2009). Collectively, this work has allowed a new way to conceptualize and assess both the stability and variability of behavior that is produced by the underlying personality system, and has opened a window into the dynamic processes within the system itself (Mischel, 2004).

Self-Control in the Cognitive–Affective Processing System

Yuichi Shoda and I proposed CAPS theory to understand how and why individuals may differ dramatically in their distinctive, stable if–then situation-behavior signatures exhibited by a given individual. CAPS is a complex system of interacting components consisting of cognitive–affective units (CAUs) that mediate between the nominal interpersonal situation the person encounters and the responses generated. [Box 26.1](#) summarizes the types of mediating CAUs hypothesized.

Some mediating CAUs encode and interpret the personal and social perceived situation in terms of the person's categories for the self, other people, and events. Some categories are chronically more accessible than others, thereby biasing social

perceptions in ways that depend on the person's social and biological history. Other CAUs represent the person's expectancies and beliefs about the self and the world, and about outcomes anticipated for behaviors in different situations. Other CAUs represent affects (feelings, emotions), values, and goals that motivate the person's plans and life projects.

Especially important for understanding self-control patterns are the CAUs representing the individual's repertoire of behavioral competencies. These are the potential behaviors that can be performed, as well as self-control strategies the person uses to regulate his or her behavior, sometimes volitionally but much more often automatically, as described above in the research on delay of gratification. These competencies include cognitive-attention strategies, plans, and scripts for generating diverse types of social behavior necessary for sustained, goal-directed effort in the pursuit of difficult goals whose attainment requires impulse control and delay of gratification (Mischel and Ayduk, 2002, 2004).

Intraindividual Variability in What's Too Hot to Handle

Intuitively, one might expect that a person who is good at self-distraction, or good at abstracting, would be good at these cognitive-attention strategies across situations, and hence there should be extremely broad cross-situational consistency in self-control, particularly since cognitive competencies and skills tend to be broader and more stable than other psychological characteristics (Mischel, 1968). In short, as common sense wisdom also suggests, some people should have more "willpower" than others, no matter what the temptations or hot trigger stimuli. And there indeed are overall aggregate differences: Some people do show overall more self-control than others and these differences are fairly stable over time, as shown in the longitudinal research on the "marshmallow [p. 15 ↓] test." But as with other characteristics, there is also impressive stable within-person if-then variability: some temptations and trigger stimuli are too hot to handle even for individuals who overall can be very effective in self-control most of the time. Recall again the Clinton example.

Box 26.1 Types of Cognitive-Affective Units in the Personality Mediating System

- 1. *Encodings*: Categories (constructs) for the self, people, events, and situations (external and internal).
- 2. *Expectancies and Beliefs*: About the social world, about outcomes for behavior in particular situations, about self-efficacy.
- 3. *Affects*: Feelings, emotions, and affective responses (including physiological reactions).
- 4. *Goals and Values*: Desirable outcomes and affective states; aversive outcomes and affective states; goals, values, and life projects.
- 5. *Competencies and Self-regulatory Plans*: Potential behaviors and scripts that one can do, and plans and strategies for organizing action and for affecting outcomes and one's own behavior and internal states.

Note: Based in part on Mischel (1973).

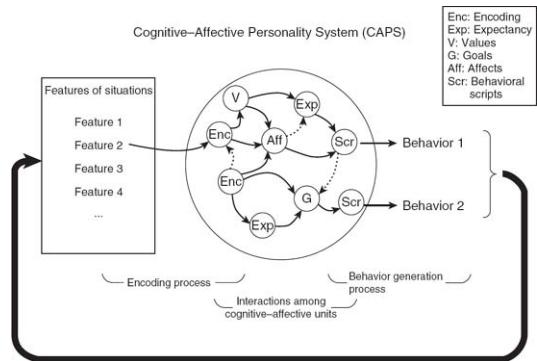
To make sense of this requires recognizing that the ease with which an individual can activate cognitive-attention strategies to cool particular hot trigger stimuli depends first of all on how hot that stimulus is for that person within the particular context. What's too hot for Clinton may not be too hot for you, and possibly vice versa. Nevertheless, reasonable consistency within and across particular domains and types of situations also can be identified (e.g., Wright and Mischel, 1987). And individual differences and within-person differences in the subjective salience and valence of different temptations are of course not the only relevant variables. Also at play are such considerations as the expectations for probable consequences that become activated, and the subjective value of those consequences, as well as their ease of activation in the particular situation. Hence no one, including any Gandhi in the world, is immune from moral dilemmas and "now" versus "later" conflicts: even people who are able and willing to cool all sorts of temptations may remain highly vulnerable to others – from addictive drugs to financial and interpersonal temptations – as even casual observers of the human condition have noted since the biblical loss of paradise.

Architecture of CAPS

The architecture of the overall CAPS system is shown in [Figure 26.2](#). Situations contain a collection of features, some of which are triggered or “turned on” by a particular nominal situation. When stimulated, these input features send “activation” into the mediating CAUs with which they are connected. The amount of activation reaching a given CA mediating unit from an activated input feature depends on importance or “weight” of that input feature's connection to that mediating unit. The aroused CAUs transmit this incoming wave of activation among themselves, ultimately settling into some internal state that will lead some response to be generated in the situation. That response in turn may change the external situation, initiating the next response cycle.

Each person is characterized by his or her own distinctive collection of cognitive-affective units with their own set of connection weights, reflecting how their learning experiences and biological histories have led [p. 16 ↓] to particular CAUs and their importance (i.e., connection weightings). The CAUs are organized into distinctive idiographic networks, broadly analogous to neural networks. Whereas each network is unique, individuals can be grouped into types and subtypes. These types may differ both on the basis of similarities in their chronic levels of accessibility (e.g., some readily access anxious expectations for rejection than others; some are more able to delay gratification, to plan effectively, and to control impulses), and on the basis of their organization (interconnections) within the system. The processing system and the situations it generates and encounters interact reciprocally and continuously in a dynamic mutual influence process.

*Figure 26.2 The Cognitive–Affective Personality System (CAPS). Situational features activate a given mediating unit which activates specific subsets of other mediating units through a stable network of relations that characterize an individual, generating a characteristic pattern of behavior in response to different situations. The relation may be positive (solid line), which increases the activation, or negative (dashed line), which decreases the activation. (From Mischel, W., and Shoda, Y. (1995). A cognitive–affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, 102, 246–268. © 1995 by the American Psychological Association. Adapted with permission.)*



Interaction of Self-Control Ability and Dispositional Vulnerability: Protective Dynamics

In the CAPS self-control model, the individual's self-regulatory ability can have important long-term protective effects that buffer against the potential negative effects of various dispositional vulnerabilities. Evidence for such a protective effective was seen in the Bing longitudinal studies discussed above. We identified these processing dynamics in studies of individual differences in the tendency to be highly rejection-sensitive (RS) in interpersonal relationships, guided by the CAPS model. Research on RS indicates that highly rejection-sensitive people in time develop lower self-esteem and become [p. 17 ↓] either more aggressive or more depressed, which, in turn, undermines the quality of their lives (Downey and Feldman, 1996). But that sequence is not inevitable. In an adult follow-up of the Bing preschoolers who had participated in the original delay of gratification studies 20 years earlier, preschool delay ability predicted adult resiliency against the potentially destructive effects of RS (Ayduk et al., 2000, Study 1). Specifically, high RS people who were able to delay gratification longer in preschool were buffered in adulthood against low self-esteem and self-worth, were better able to cope with stress, and had greater ego resiliency. High RS people who were unable to delay gratification in preschool had lower academic achievement and more frequent

cocaine/crack use than low RS people. In contrast, high RS people who had high preschool delay ability were buffered against such negative outcomes.

In CAPS self-control theory, these findings reflect differences in people's ability to readily (automatically) activate strategies for "cooling" and reducing the "hot thoughts" to which their RS makes them vulnerable. Thereby they can avoid the impulsive reactions (e.g., become enraged, create a fight) to which they are vulnerable. A parallel study, conducted among low-income, urban, minority middle-school children who are at higher demographic risk for maladjustment, replicated these findings with population-appropriate measures (Ayduk et al., 2000, Study 2). Again, among children high in RS, delay of gratification ability was associated with lower aggression against peers, greater interpersonal acceptance, and higher levels of self-worth. Children who were low in delay ability but high in RS exhibited the negative behaviors typical of the RS dynamics. The overall findings support the idea that self-regulatory competencies, and the cooling mechanisms they enable, restrain the negative influence of high RS on the behavior that plays out. The seeds of individual differences in attention control and self-regulation are visible already in the behavior of toddlers in their ability to cool their negative affect during brief separation from the mother. They in turn predict self-regulatory competencies years later, as shown in more adaptive patterns of "cool" attention control when they deal with the challenges of the marshmallow test at age five years (Sethi et al., 2000).

Agency in CAPS

CAPS is not a passive, reactive system: it is agentic and proactive in the sense that it also acts upon itself through a feedback loop, both by generating its own internal situations (e.g., in anticipated and planned events, in fantasy, in self-reflection), and through the behaviors that the system generates in interaction with the social world. Such behaviors (e.g., impulsive reactions, failures to carry out intentions, effective control efforts and goal pursuit) further influence the individual's social–cognitive experiences and evolving social learning history, and modify the subsequent situations encountered and generated. In this view, development of the self-regulatory system is a lifelong process of adaptation both through assimilating new stimuli into the existing CAPS network and by accommodating the network itself in response to novel situations.

Prospects

The questions that motivated my work beginning around the kitchen table talking with my three little daughters 40 years ago have yielded valuable answers, sure to change in the future, but pointing to core cognitive-affective mechanisms that help to demystify “willpower.” We are eager to probe ever more deeply into the basic underlying mechanisms that control delay of gratification behavior and impulse control, turning now to the neural and brain level of analysis. Therefore we have formed an interdisciplinary team [p. 18 ↓] with cognitive neuroscientists, including B.J. Casey, John Jonides, Ozlem Ayduk, Kevin Ochsner, Edward E. Smith, Yuichi Shoda, and other colleagues to uncover with increasing depth the neural as well as the cognitive, affective, and social-behavioral mechanisms that enable impulse control. Participants in the Bing Longitudinal Study have been invited to the Lucas Center for Imaging at Stanford University but only a few have been scanned to date (Mischel et al., in press). As the participants are reaching middle adulthood, we also are continuing to assess consequential outcomes, including occupational and marital status, economic behavior, social, cognitive, and emotional functioning, as well as mental and physical health and wellbeing. The results we trust will help explain with increasing precision the individual differences in “willpower” revealed in the longitudinal studies, perhaps even extending into the final years of life.

So What?

Not just psychologists, but also economists, policy makers, and educators (as well as the media) have now recognized that the “marshmallow test” taps important long-term individual differences in self-control early in life that predict highly consequential mental and physical health consequences over much of the lifespan, and do so far better than intelligence tests. That kind of long-term predictability for important life outcomes is remarkably rare, if not unique, within psychological science. Most exciting to me, beyond prediction, and of particular social significance, the findings to date help reveal how the mechanisms that underlie willpower can allow people who have difficulty with delay of gratification (including the author) to do so. They can learn strategies to

control their attention strategically and to change how the temptations are mentally represented to “cool” their impact. The power of such cognitive reappraisal has been amply demonstrated, at least for short-term changes in the laboratory. The challenge now is to see how these strategies can best be taught early in life and maintained virtually automatically for long-term enhancement of self-control ability. The implications for educational and social policy and for therapeutic interventions are potentially enormous, as is now widely recognized even in the media, for example in a *New York Times* editorial column by David Brooks entitled “Marshmallows and Public Policy” (e.g., Brooks, 2006; Gladwell, 2002; Goleman, 2006; Lehrer, 2009).

With the goal of creating educational interventions designed to enhance the psychological skills underlying willpower phenomena, Angela Duckworth and a team of colleagues and I are currently pilot testing such attempts within the schools. The current policies and practices of the American school system, beginning in the early grades, are primarily designed to cultivate knowledge and analytical skills. But it has become evident that such skills are not the only, nor even the most important, competencies essential to successful functioning and flourishing in adulthood. Our science is now poised to identify and enhance with increasing precision the psychological skills and strategies that enable “willpower” in young children so that they can learn to use them to realize their full cognitive and social potential. These skills constitute what Goleman (2006) has labeled the “master aptitude.” They promise to be teachable, and with surprisingly simple theory-based core strategies (e.g., Kross et al., 2010; Mischel and Ayduk, 2004). If willpower can be dramatically facilitated by teaching learnable skills and strategies to the young child, the educational policy and therapeutic implications are as evident as they are profound.

References

Adolphs, R., Tranel, D., Hamann, S., Young, A.W., Calder, A.J., Phelps, E.A., Anderson, A., Lee, G.P. and Damasio, A.R. Recognition of facial emotion in nine individuals with bilateral amygdala damage . *Neuropsychologia* , 37, 1111–1117. (1999)
<http://dx.doi.org/10.1016/S0028-3932%2899%2900039-1>

Ainslie, G. (2001) Breakdown of Will . Cambridge: Cambridge University Press.

Andersen, S.M. and Chen, S. The relational self: An interpersonal social-cognitive theory . *Psychological Review* , 109, 619–645. (2002) <http://dx.doi.org/10.1037/0033-295X.109.4.619>

Ayduk, O., Mendoza-Denton, R., Mischel, W., Downey, G., Peake, P. and Rodriguez, M.L. Regulating the interpersonal self: Strategic self-regulation for coping with rejection sensitivity . *Journal of Personality and Social Psychology* , 79, 776–792. (2000) <http://dx.doi.org/10.1037/0022-3514.79.5.776>

Ayduk, O. and Mischel, W. (2002) When smart people behave stupidly: Inconsistencies in social and emotional intelligence . In R.J. Sternberg (ed.), *Why Smart People Can Be So Stupid* , pp. 86–105. New Haven, CT: Yale University Press.

Ayduk, O., Zayas, V., Downey, G., Cole, A.B., Shoda, Y. and Mischel, W. Rejection sensitivity and executive control: Joint predictors of borderline personality features . *Journal of Research in Personality* , 42, 151–168. (2008) <http://dx.doi.org/10.1016/j.jrp.2007.04.002>

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory* . Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. and Mischel, W. Modification of self-imposed delay of reward through exposure to live and symbolic models . *Journal of Personality and Social Psychology* , 2, 698–705. (1965) <http://dx.doi.org/10.1037/h0022655>

Berlyne, D. (1960) Conflict, Arousal, and Curiosity . New York: McGraw Hill. <http://dx.doi.org/10.1037/11164-000>

Borkenau, P., Riemann, R., Spinath, F.M. and Angleitner, A. Genetic and environmental influences on person situation profiles . *Journal of Personality* , 74, 1451–1479. (2006) <http://dx.doi.org/10.1111/j.1467-6494.2006.00416.x>

Brooks, D. (2006) Marshmallows and public policy . *The New York Times* , 7 May.

Cervone, D., ed. and Shoda, Y. (eds) (1999) *The Coherence of Personality: Social-Cognitive Bases of Consistency, Variability, and Organization*. New York: Guilford Press.

Downey, G. and Feldman, S. Implications of rejection sensitivity for intimate relationships . *Journal of Personality and Social Psychology* , 70, 1327–1343. (1996) <http://dx.doi.org/10.1037/0022-3514.70.6.1327>

Eigsti, I., Zayas, V., Mischel, W., Shoda, Y., Ayduk, O., Dadlani, M.B., Davidson, M.C., Aber, J.L. and Casey, B.J. Predicting cognitive control from preschool to late adolescence and young adulthood . *Psychological Science* , 17, 478–484. (2006) <http://dx.doi.org/10.1111/j.1467-9280.2006.01732.x>

Eisenberg, N., Fabes, R.A., Guthrie, I.K. and Reiser, M. (2002) The role of emotionality and regulation in children's social competence and adjustment . In L. Pulkkinen, ed. and A. Caspi (eds), *Paths to Successful Development: Personality in the Life Course* , pp. 46–70. New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511489761.003>

Eisenberger, N., Smith, C.L., Sadovsky, A. and Spinrad, T.L. (2004) Effortful control: reactions with emotion regulation, adjustment, and socialization in childhood . In R.F. Baumeister, ed. and K.D. Vohs (eds), *Handbook of Self Regulation* . New York: Guilford Press.

Epstein, S. Integration of the cognitive and psychodynamic unconscious . *American Psychologist* , 49, 709–724. (1994) <http://dx.doi.org/10.1037/0003-066X.49.8.709>

Estes, W.K. Reinforcement in human behavior . *American Scientist* , 60, 723–729. (1972)

Fournier, M.A., Moskowitz, D.S. and Zuroff, D.C. Integrating dispositions, signatures, and the interpersonal domain . *Journal of Personality and Social Psychology* , 94, 531–545 754–768. (2008) <http://dx.doi.org/10.1037/0022-3514.94.3.531>

Freud, S. (1959) Formulations regarding the two principles of mental functioning . Collected Papers, Vol. IV . New York: Basic Books. (Originally published 1911).

Gladwell, M. (2002) *The Tipping Point: How Little Things Can Make a Big Difference*. New York: Little Brown and Company.

Goleman, D. (2006) *Emotional Intelligence: Why It Can Matter More Than IQ*. New York: Bantam.

Gollwitzer, P.M. Implementation intentions: Strong effects of simple plans. *American Psychologist*, 54, 493–503. (1999) <http://dx.doi.org/10.1037/0003-066X.54.7.493>

Gray, J.A. (1982) *The Neuropsychology of Anxiety*. Oxford: Oxford University Press.

Gray, J.A. (1987) *The Psychology of Fear and Stress*, 2nd Edition. New York: McGraw-Hill.

Hull, C.L. Goal attraction and directing ideas conceived as habit phenomena. *Psychological Review*, 38, 487–506. (1931) <http://dx.doi.org/10.1037/h0071442>

Johnson, M.H., Posner, M.I. and Rothbart, M.K. Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. *Journal of Cognitive Neuroscience*, 3, 335–344. (1991) <http://dx.doi.org/10.1162/jocn.1991.3.4.335>

Jackson, D.C., Muller, C.J., Dolski, I., Dalton, K.M., Nitschke, J.B., Urry, H.L., Rosenkranz, M.A., Ryff, D.C., Singer, B.H. and Davidson, R.J. Now you feel it, now you don't: Frontal brain electrical asymmetry and individual differences in emotion regulation. *Psychological Science*, 14, 612–617. (2003) http://dx.doi.org/10.1046/j.0956-7976.2003.psci_1473.x

Kross, E., Mischel, W. and Shoda, Y. (2010) Enabling self-control: A cognitive affective processing system (CAPS) approach to problematic behavior. In J. Maddux, ed. and J. Tangney (eds), *Social Psychological Foundations of Clinical Psychology*, pp. 375–394. New York: Guilford.

Kross, E. and Ochsner, K. (2010) Integrating research on self-control across multiple levels of analysis: A social cognitive neuroscience approach. In R. Hassin, ed. , K. Ochsner, ed. and Y. Trope. (eds), *From Society to Brain: The New Sciences of Self-*

Control , pp. 76–92. New York: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780195391381.003.0005>

LeDoux, J. (1996) The Emotional Brain . New York: Touchstone.

LeDoux, J.E. Emotion circuits in the brain . Annual Review of Neuroscience , 23, 155–184. (2000) <http://dx.doi.org/10.1146/annurev.neuro.23.1.155>

Lehrer, J. (2009) Don't! The secret of self-control . The New Yorker , 18 May.

Lieberman, M.D. Social cognitive neuroscience: A review of core processes . Annual Review of Psychology , 58, 259–89. (2007) <http://dx.doi.org/10.1146/annurev.psych.58.110405.085654>

Lieberman, M.D., Gaunt, R., Gilbert, D.T. and Trope, Y. (2002). Reflection and reflexion: A social cognitive neuroscience approach to attributional inference . In M. Zanna (ed.), Advances in Experimental Social Psychology , 34, 199–249. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2802%2980006-5>

Loewenstein, G., ed. , Read, D., ed. and Baumeister, R. (eds) (2003) Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice . New York: Russell Sage Foundation.

Metcalfe, J. and Jacobs, W.J. A 'hot/cool-system' view of memory under stress . PTSD Research Quarterly , 7, 1–6. (1996)

Metcalfe, J. and Jacobs, W.J. (1998) Emotional memory: The effects of stress on 'cool' and 'hot' memory systems . In D.L. Medin (ed.), The Psychology of Learning and Motivation: Advances in Research and Theory , 38, 187–222. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0079-7421%2808%2960187-5>

Metcalfe, J. and Mischel, W. A hot/cool system analysis of delay of gratification: Dynamics of willpower . Psychological Review , 106, 3–19. (1999) <http://dx.doi.org/10.1037/0033-295X.106.1.3>

Mischel, W. Preference for delayed reinforcement and social responsibility . *Journal of Abnormal and Social Psychology* , 62, 1–7. (1961a) <http://dx.doi.org/10.1037/h0048263>

Mischel, W. Delay of gratification, need for achievement and acquiescence in another culture . *Journal of Abnormal and Social Psychology* , 62, 543–552. (1961b) <http://dx.doi.org/10.1037/h0039842>

Mischel, W. Father absence and delay of gratification: Cross-cultural comparisons . *Journal of Abnormal and Social Psychology* , 63, 116–124. (1961c) <http://dx.doi.org/10.1037/h0046877>

Mischel, W. (1966) Theory and research on the antecedents of self-imposed delay of reward . In B.A. Maher (ed.), *Progress in Experimental Personality Research* , 3, 85–131. New York: Academic Press.

Mischel, W. (1968) *Personality and Assessment* . New Jersey: Erlbaum.

Mischel, W. Toward a cognitive social learning reconceptualization of personality . *Psychological Review* , 80, 252–283. (1973) <http://dx.doi.org/10.1037/h0035002>

Mischel, W. (1974). Processes in delay of gratification . In L. Berkowitz (ed.), *Advances in Experimental Social Psychology* , 7, 249–292. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960039-8>

Mischel, W. Toward an integrative science of the person . *Annual Review of Psychology* , 55, 1–22. (2004) <http://dx.doi.org/10.1146/annurev.psych.55.042902.130709>

Mischel, W. From Personality and Assessment to personality science . *Journal of Research in Personality (Special Issue: Personality and Assessment 40 years later)* , 43, 282–290. (1968) (2009)

Mischel, W. and Ayduk, O. Self-regulation in a cognitive-affective personality system: Attentional control in the service of the self . *Self and Identity* , 1, 113–120. (2002) <http://dx.doi.org/10.1080/152988602317319285>

Mischel, W. and Ayduk, O. (2004) Willpower in a cognitive-affective processing system: The dynamics of delay of gratification . In R.F. Baumeister, ed. and K.D. Vohs (eds), *Handbook of Self-regulation: Research, Theory, and Applications* , pp. 99–129. New York: Guilford Press.

Mischel, W., Ayduk, O., Berman, M., Casey, B.J., Gotlib, I., Jonides, J., Kross, E., Teslovich, T., Wilson, N., Zayas, V. and Shoda, Y. (in press) 'Willpower' over the life span: decomposing impulse control . *Social and Cognitive Affective Neuroscience* (Special Issue on Social and Affective Neuroscience and Neuroeconomics of Aging) .

Mischel, W. and Baker, N. Cognitive appraisals and transformations in delay behavior . *Journal of Personality and Social Psychology* , 31, 254–261. (1975) <http://dx.doi.org/10.1037/h0076272>

Mischel, W. and Ebbesen, E.B. Attention in delay of gratification . *Journal of Personality and Social Psychology* , 16, 239–337. (1970) <http://dx.doi.org/10.1037/h0029815>

Mischel, W., Ebbesen, E.B. and Zeiss, A.R. Cognitive and attentional mechanisms in delay of gratification . *Journal of Personality and Social Psychology* , 21, 204–218. (1972) <http://dx.doi.org/10.1037/h0032198>

Mischel, W. and Gilligan, C. Delay of gratification, motivation for the prohibited gratification, and resistance to temptation . *Journal of Abnormal and Social Psychology* , 69, 411–417. (1964) <http://dx.doi.org/10.1037/h0048918>

Mischel, W. and Metzner, R. Preference for delayed reward as a function of age, intelligence, and length of delay interval . *Journal of Abnormal and Social Psychology* , 64, 425–431. (1962) <http://dx.doi.org/10.1037/h0045046>

Mischel, W. and Mischel, F. Psychological aspects of spirit possession: A reinforcement analysis . *American Anthropologist* , 60, 249–260. (1958) <http://dx.doi.org/10.1525/aa.1958.60.2.02a00040>

Mischel, W. and Moore, B. Effects of attention to symbolically-presented rewards on self-control . *Journal of Personality and Social Psychology* , 28, 172–179. (1973) <http://dx.doi.org/10.1037/h0035716>

Mischel, W. and Patterson, C.J. Substantive and structural elements of effective plans for self-control . *Journal of Personality and Social Psychology* , 34, 942–950. (1976) <http://dx.doi.org/10.1037/0022-3514.34.5.942>

Mischel, W. and Peake, P.K. Beyond deja vu in the search for cross-situational consistency . *Psychological Review* , 89, 730–755. (1982a) <http://dx.doi.org/10.1037/0033-295X.89.6.730>

Mischel, W. and Peake, P. (1982b) In search of consistency: Measure for measure . In M.P. Zanna, ed. , E.T. Higgins, ed. , and C.P. Herman (eds), *Consistency in Social Behavior: The Ontario Symposium* , 2, 187–207. Hillsdale, NJ: Erlbaum.

Mischel, W. and Shoda, Y. A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure . *Psychological Review* , 102, 246–268. (1995) <http://dx.doi.org/10.1037/0033-295X.102.2.246>

Mischel, W., Shoda, Y. and Peake, P.K. The nature of adolescent competencies predicted by preschool delay of gratification . *Journal of Personality and Social Psychology* , 54, 687–699. (1988) <http://dx.doi.org/10.1037/0022-3514.54.4.687>

Mischel, W., Shoda, Y. and Rodriguez, M.L. Delay of gratification in children . *Science* , 244, 933–938. (1989) <http://dx.doi.org/10.1126/science.2658056>

Mischel, W. and Staub, E. Effects of expectancy on working and waiting for larger rewards . *Journal of Personality and Social Psychology* , 2, 625–633. (1965) <http://dx.doi.org/10.1037/h0022677>

Moore, B., Mischel, W. and Zeiss, A. Comparative effects of the reward stimulus and its cognitive representation in voluntary delay . *Journal of Personality and Social Psychology* , 34, 419–424. (1976) <http://dx.doi.org/10.1037/0022-3514.34.3.419>

Morf, C.C. and Rhodewalt, F. Unraveling the paradoxes of narcissism: A dynamic self-regulatory processing model . *Psychological Inquiry* , 12, 177–196. (2001) http://dx.doi.org/10.1207/S15327965PLI1204_1

Morrow, L. (1998, Feb. 2) The reckless and the stupid: A character of a President eventually determines his destiny . Time Magazine .

Moskowitz, D.S., Suh, E.J. and Desaulniers, J. Situational influences on gender differences in agency and communion . Journal of Personality and Social Psychology , 66, 753–761. (1994) <http://dx.doi.org/10.1037/0022-3514.66.4.753>

Nolen-Hoeksema, S. Responses to depression and their effects on the duration of depressive episodes . Journal of Abnormal Psychology , 100, 569–582. (1991) <http://dx.doi.org/10.1037/0021-843X.100.4.569>

Ochsner, K.N. and Gross, J.J. The cognitive control of emotion . Trends in Cognitive Science , 27, 26–36. (2005)

Ochsner, K.N. and Gross, J.J. (2007) The neural architecture of emotion regulation . In J.J. Gross, ed. and R. Buck (eds), The Handbook of Emotion Regulation , pp. 87–109. New York: Guilford Press.

Ochsner, K.N. and Gross, J.J. Cognitive emotion regulation: Insights from social cognitive and affective neuroscience . Currents Directions in Psychological Science , 17, 153–158. (2008) <http://dx.doi.org/10.1111/j.1467-8721.2008.00566.x>

Orom, H. and Cervone, D. Personality dynamics, meaning, and idiosyncrasy: Identifying cross-situational coherence by assessing personality architecture . Journal of Research in Personality (Special Issue: Personality and Assessment 40 years later) , 43, 228–240. (2009)

Patterson, C.J. and Mischel, W. Effects of temptation-inhibiting and task-facilitating plans on self-control . Journal of Personality and Social Psychology , 33, 209–217. (1976) <http://dx.doi.org/10.1037/0022-3514.33.2.209>

Peake, P., Hebl, M. and Mischel, W. Strategic attention deployment in waiting and working situations . Developmental Psychology , 38, 313–326. (2002) <http://dx.doi.org/10.1037/0012-1649.38.2.313>

- Phelps, E.A., O'Connor, K.J., Gatenby, J.C., Gore, J.C., Grillon, C. and Davis, M. Activation of the left amygdala to a cognitive representation of fear . *Nature Neuroscience* , 4, 437–441. (2001) <http://dx.doi.org/10.1038/86110>
- Posner, M.I., Rothbart, M.K., Gerardi, G. and Thomas-Thrapp, L.J. (1997) Functions of orienting in early infancy . In P. Lang, ed. , M. Balaban, ed. , and R.F. Simmons (eds), *The Study of Attention: Cognitive Perspectives from Psychophysiology, Reflexology and Neuroscience* , pp. 327–345. Hillsdale, NJ: Erlbaum.
- Rachlin, H. (2000) *The Science of Self-control* . Cambridge, MA: Harvard University Press.
- Rapaport, D. (1967) *The Collected Papers of David Rapaport* . New York: Basic Books.
- Rodriguez, M.L., Mischel, W. and Shoda, Y. Cognitive person variables in the delay of gratification of older children at-risk . *Journal of Personality and Social Psychology* , 57, 358–367. (1989) <http://dx.doi.org/10.1037/0022-3514.57.2.358>
- Rothbart, M.K., Ellis, L.K. and Posner, M.I., (2004) Temperament and self-regulation . In R.F. Baumeister, ed. and K.D. Vohs (eds), *Handbook of Self-Regulation* . New York: Guilford Press.
- Rusting, C.L. and Nolen-Hoeksema, S. Regulating responses to anger: Effects on rumination and distraction on angry mood . *Journal of Personality and Social Psychology* , 74, 790–803. (1998) <http://dx.doi.org/10.1037/0022-3514.74.3.790>
- Sethi, A., Mischel, W., Aber, L., Shoda, Y. and Rodriguez, M. The role of strategic attention deployment in development of self-regulation: Predicting preschoolers' delay of gratification from mother-toddler interactions . *Developmental Psychology* , 36, 767–777. (2000) <http://dx.doi.org/10.1037/0012-1649.36.6.767>
- Shoda, Y. and LeeTiernan, S.J. (2002) What remains invariant? Finding order within a person's thoughts, feelings, and behaviors across situations . In D. Cervone, ed. and W. Mischel (eds), *Advances in Personality Science* , pp. 241–270. New York: Guilford Press.

Shoda, Y., Mischel, W. and Peake, P.K. Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: Identifying diagnostic conditions . *Developmental Psychology* , 26, 978–986. (1990) <http://dx.doi.org/10.1037/0012-1649.26.6.978>

Shoda, Y., Mischel, W. and Wright, J. The role of situational demands and cognitive competencies in behavior organization and personality coherence . *Journal of Personality and Social Psychology* , 65, 1023–1035. (1993) <http://dx.doi.org/10.1037/0022-3514.65.5.1023>

Shoda, Y., Mischel, W. and Wright, J.C. Intra-individual stability in the organization and patterning of behavior: Incorporating psychological situations into the idiographic analysis of personality . *Journal of Personality and Social Psychology* , 67, 674–687. (1994) <http://dx.doi.org/10.1037/0022-3514.67.4.674>

Trope, Y. and Liberman, N. Temporal construal . *Psychological Review* , 110, 403–421. (2003) <http://dx.doi.org/10.1037/0033-295X.110.3.403>

Van Mechelen, I. A royal road to understanding the mechanisms underlying person-in-context behavior . *Journal of Research in Personality (Special Issue: Personality and Assessment 40 years later)* , 43, 179–186. (2009)

Vansteelandt, K. and Van Mechelen, I. Individual differences in situation-behavior profiles: A triple typology model . *Journal of Personality and Social Psychology* , 75, 751–765. (1998) <http://dx.doi.org/10.1037/0022-3514.75.3.751>

Winston, J.S., Strange, B.A., O'Doherty, J. and Dolan, R.J. Automatic and intentional brain responses during evaluation of trustworthiness of faces . *Nature Neuroscience* , 5, 277–283. (2002) <http://dx.doi.org/10.1038/nn816>

Wright, J.C. and Mischel, W. A conditional approach to dispositional constructs: The local predictability of social behavior . *Journal of Personality and Social Psychology* , 53, 1159–1177. (1987) <http://dx.doi.org/10.1037/0022-3514.53.6.1159>

Wright, J.C. and Mischel, W. Conditional hedges and the intuitive psychology of traits . Journal of Personality and Social Psychology , 55, 454–469. (1988) <http://dx.doi.org/10.1037/0022-3514.55.3.454>

<http://dx.doi.org/10.4135/9781446249222.n26>

Handbook of Theories of Social Psychology

Self-Verification Theory

Contributors: William B. Swann Jr.

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Self-Verification Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n27>

Print pages: 23-43

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n27>

[p. 23 ↓]

Chapter 27: Self-Verification Theory

William B. Swann, Jr.

Abstract

Self-verification theory proposes that people prefer others to see them as they see themselves, even if their self-views happen to be negative. For example, those who see themselves as likable want others to see them as such, and people who see themselves as dislikable want others to perceive them that way. Presumably, people seek self-verification because self-verifying evaluations make the world seem coherent and predictable. In addition, self-verifying evaluations smooth social interaction by guiding action and letting people know what to expect from others. People strive for self-verification by gravitating toward interaction partners and settings that seem likely to provide self-confirming evaluations. Moreover, once in relationships, people actively evoke self-confirming reactions from their partners. Finally, people process feedback about themselves in ways that promote the survival of their self-views. In general, self-verification strivings are adaptive and functional, as they foster feelings of coherence, reduce anxiety, improve group functioning, and erode social stereotypes. Nevertheless, for those who possess inappropriately negative self-views, self-verification may thwart positive change and make their life situations harsher than they would be otherwise. In this chapter, I discuss the nature, history, and social implications of self-verification theory and research.

Introduction

It all started with a seven year old boy named Tommy. I met Tommy while working at a camp for underprivileged children following my sophomore year in college. I still have a vivid memory of our first encounter. It was my first day at the camp, and I was eager

to meet the kids. As I approached the camp director's cabin, however, I was alarmed at the sound of some boys fighting. I ran over to find Tommy on the ground, pinned down by two other children who were wailing on him mercilessly. A couple of other adults (counselors, I learned later) and I stepped in to break up the fight. Someone escorted Tommy to the nurse's office to repair the damage, which was minor.

This was the first of my many memorable encounters with Tommy. Unfortunately, these encounters were rarely happy occasions. As the camp director sadly noted, Tommy was a little cloud that hung over "Camp Sunshine," reigning difficulties on almost everyone he encountered. The director then noted that my application indicated that I was a psychology [p. 24 ↓] major, which led her to wonder if I might be interested in trying to figure out what was the matter with Tommy. I hesitated before answering. At this point in my life I did not suffer from lack of confidence, but I had enough humility to recognize that there was little hope that I could develop a deep understanding of a character as complex as Tommy, especially in the span of a few months. Nevertheless, I was fascinated by the young boy and his seemingly bizarre behavior. Intrigued, I agreed to spend some time observing Tommy and report back to the director.

Over the next few weeks my fascination with Tommy grew, for I was completely unprepared for what I observed. In his interactions, Tommy seemed hell bent on turning everyone against him: disobeying the counselors, taunting and teasing the other kids, and being generally disruptive. His relationship with "Crazy Louis" was particularly remarkable. Louis earned his "Crazy" label by ruthlessly assaulting the other children on a daily basis. Often his aggressiveness seemed random and unprovoked. All of the children rapidly learned to steer clear of Louis – except for Tommy, that is. Tommy seemed drawn to Louis like a magnet. Louis would oblige by subjecting Tommy to a steady diet of verbal and physical abuse.

And each evening, when I talked to Tommy about his day, he remembered only the negatives – the problems he encountered and the slights that had been directed at him. In contrast, when I mentioned the positive things that had happened he seemed confused, forgetful, and anxious, returning as quickly as possible to his narrative of negativity.

What puzzled me about Tommy was that his activities seemed almost tailored made to sour his relations with others and perpetuate his incredibly negative self-image. When I probed, it seemed like he derived some comfort from the fact that his experiences at the camp were every bit as bad as he expected them to be. Tommy not only seemed convinced that the world hated him; he seemed reassured when his interactions supported this expectation.

Tommy's pathology became easier to understand after I consulted the case worker who referred Tommy to the camp. She revealed that he had been the target of a steady stream of abuse since he was an infant. Apparently, he had internalized the treatment he received. An incredibly negative identity resulted. It was not surprising to me that Tommy's negative self-views could be traced to terrible experiences with his caregivers. What was surprising was that he seemed to work actively to recreate the negative conditions that generated his negative identity in the first place. Most people would seemingly want to escape an ugly past rather than recreate it. What made Tommy different?

It would take me years before I would get a handle on this question, for as an undergraduate I lacked the sophistication to address it in a meaningful way. My efforts to acquire the training I needed jumpstarted when I gained admission to graduate school in social psychology. From my home in Pennsylvania, I headed north to the University of Minnesota. There I began working with Mark Snyder, an eminent scholar with interests in the self and social interaction. When I arrived I learned that he was about to launch an exciting new program of research. The topic was the self-fulfilling effects of the expectations of some persons ("perceivers") on the behaviors of their interaction partners ("targets"). This phenomenon seemed to represent the flip side of the activities of Tommy, a "target" whose self-views influenced the behavior of all of the "perceivers" around him. I happily immersed myself in this project, and was later rewarded with three publications (Snyder and Swann, 1978a, 1978b; Swann and Snyder, 1980).

It was not until my final year at Minnesota that Tommy reappeared on my intellectual radar screen. In designing my dissertation, I decided to test the relative power of the expectations of perceivers and the self-views of targets. Guided by my experiences with [p. 25 ↓] Tommy, I expected that targets who had firmly held self-views would

repudiate expectations that challenged their self-views, even if their self-views were negative. This was precisely what happened – people with negative self-views elicited more negative reactions than people with positive self-views. Moreover, the tendency for participants to elicit negative self-confirming reactions was particularly strong when they suspected that their interaction partner held positive appraisals of them.

Upon completion of my dissertation, I took a job at the University of Texas at Austin. There, I conducted several follow-ups to my dissertation research with Stephen Read. Those studies were packaged together in two papers that appeared in years to follow (Swann and Read, 1981a, 1981b). The core argument that Steve and I advanced was that people were like Tommy in that they wanted to confirm their self-views. We also suggested that they expressed this preference during each of three successive phases of the interaction sequence. In Study 1, we examined attention. We recruited participants who perceived themselves as either likable or dislikable and told them that another person had likely evaluated them in either a positive or negative manner. The question was how long participants would read a passage that they (erroneously) thought that the evaluator had written about them. Participants who saw themselves as likable spent longer reading the passage when they expected it to be positive. In contrast, those with negative self-views spent longer reading the passage when they expected it to be negative. Study 2, my dissertation study, showed that people behaved in ways that elicited reactions from their interaction partners that confirmed their self-views. Study 3 focused on what participants remembered about evaluations they received. We discovered that participants preferentially recalled self-verifying evaluations. These data offered compelling support for our hypotheses: within each of three distinct phases of social interaction, people sought to verify their self-views.

In a series of follow-up studies, we tested the notion that people seek and value self-verifying evaluations because such evaluations more informative and diagnostic than nonverifying evaluations. Participants in Study 1 preferentially solicited feedback that verified their self-views, whether these self-views were positive or negative. In Study 2, participants spent more money to obtain verifying as compared to nonverifying evaluations. Study 3 revealed that participants perceived self-verifying evaluations to be particularly informative and diagnostic.

Together, the results presented in the Swann and Read papers strongly suggested that Tommy was no anomaly. Rather, there seemed to be a fairly robust tendency for people to prefer self-confirming feedback over nonconfirming feedback. In fact, this preference influenced information seeking, attention, memory, overt behavior, and even perceptions of the diagnosticity of the feedback. These studies provided the empirical foundation on which the theory would rest. The next task was to flesh out the theory and begin to explore its implications. My efforts culminated in the publication of a chapter in which I presented the essential elements of this theory (Swann, 1983).

Self-Verification Theory

The core idea underlying self-verification theory was first articulated by Prescott Lecky (1945). He proposed that chronic self-views give people a strong sense of coherence and they are thus motivated to maintain them. Related ideas resurfaced a few years later in several self-consistency theories (e.g., Aronson, 1968; Festinger, 1957; Secord and Backman, 1965). Nevertheless, the most prominent consistency theorists transformed Lecky's theory in a fundamental way, for the emphasis on experimentation during that era led to the abandonment of Lecky's emphasis on the role of chronic self-views in [p. 26 ↓] consistency strivings. Dissonance theory (Aronson, 1968; Festinger, 1957), for example, emphasized the ways in which people found consistency by bringing their transient self-images into accord with their overt behaviors. Self-verification theory (Swann, 1983) reversed this trend by reinstating Lecky's belief that stable self-views organize people's efforts to maximize consistency. Therefore, rather than changing self-views willy nilly to match behavior, self-verification theory holds that people are motivated to maximize the extent to which their experiences confirm and reinforce their self-views.

People's powerful allegiance to stable self-views can be understood by considering how and why they develop self-views in the first place. Theorists have long assumed that people form their self-views by observing how others treat them (e.g., Cooley, 1902; Mead, 1934). As they acquire more and more evidence to support their self-views, people become increasingly certain of them. When certainty increases enough, people begin using their self-views in making predictions about their worlds, guiding behavior, and maintaining a sense of coherence, place, and continuity. In this way, stable self-

views not only serve the pragmatic function of guiding behavior, they also serve the epistemic function of affirming people's sense that things are as they should be. Indeed, firmly held self-views form the centerpiece of their knowledge systems. As such, when people strive for self-verification, the viability of that system hangs in the balance. It is thus not surprising that by mid childhood, a preference for evaluations that confirm and stabilize self-views emerges (e.g., Cassidy et al., 2003).

The origins of the self-verification motive can also be understood from an evolutionary perspective. Evolutionary biologists generally agree that humans spent most of their evolutionary history in small hunter-gatherer groups. Self-verification strivings would have been advantageous in such groups. That is, once people used inputs from the social environment to form self-views, self-verification strivings would have stabilized their identities and behavior, which in turn would make each individual more predictable to other group members (e.g., Goffman, 1959). Mutual predictability would facilitate division of labor, making the group more effective in accomplishing its objectives. Ultimately, the stable self-views fostered by self-verification strivings would bolster survival rates of group members (see Leary and Baumeister's [2000] sociometer theory for another perspective on the utility of accurate self-knowledge for group functioning).

The desire for stable self-views produced by self-verification strivings may also be understood on a neurological level. Of their very nature, self-verifying evaluations will be more predictable and familiar than non-verifying ones. Such stimuli are not only more "perceptually fluent" (more readily processed) than unpredictable and unfamiliar stimuli, they have also been shown to foster positive affect (e.g., Winkielman et al., 2002). The preference for self-verifying evaluations may therefore stem, at least partially, from basic properties of the human brain.

If stable self-views are essential to human functioning, those who are deprived of them should be seriously impaired. This seems to be true. Witness a case study reported by the neurologist Oliver Sacks (1985). Due to chronic alcohol abuse, patient William Thompson suffered from memory loss that was so profound that he forgot who he was. Only able to remember scattered fragments from his past, Thompson lapsed into a state of psychological anarchy. But Thompson did not give up. Instead, he desperately attempted to recover the self that eluded him. For instance, he sometimes developed hypotheses about who he was and then tested these hypotheses on

whoever happened to be present. For example, thinking he was a customer at a butcher shop, he approached another patient and tried to identify him: “You must be Hymie, the Kosher butcher next door … But why are there no bloodstains on your coat?” Tragically, Thompson [p. 27 ↓] could never remember the results of his latest “test.” He was thus doomed to enact such tests repeatedly for the remainder of his life.

Thompson’s case not only shows that stable self-views are essential to psychological wellbeing, it also shows how essential such self-views are to guiding action. Plagued by a sense of self that kept disappearing like the Cheshire Cat, Thompson did not know how to act toward people. In a very real sense, his inability to obtain self-verification deprived him of his capacity to have meaningful interactions with the people around him. No wonder, then, that people enact numerous strategies designed to elicit support for their self-views.

How Self-Verification Strivings Shape Social Reality

People may use three distinct processes to create self-verifying social worlds. First, people may construct self-verifying “opportunity structures;” that is, social environments that satisfy their needs (McCall and Simmons, 1966). They may, for example, seek and enter relationships in which they are apt to enjoy confirmation of their self-views (e.g., Swann et al., 1989) and leave relationships in which they fail to receive self-verification (Swann et al., 1994).

A second self-verification strategy involves the systematic communication of self-views to others. For example, people may display “identity cues” – highly visible signs and symbols of who they are. Physical appearances represent a particularly potent class of identity cues. The clothes one wears, for instance, can advertise numerous self-views, including one’s political leanings, income level, religious convictions, and so on (e.g., Gosling, 2008; Pratt and Rafaeli, 1997). Even email addresses can communicate identities to others (Chang-Schneider and Swann, 2009).

People may also communicate their identities to others through their actions. Depressed college students, for example, were more likely to solicit unfavorable feedback from their roommates than were nondepressed students (Swann et al., 1992d). Such efforts bore fruit in the form of negative evaluations. That is, the more unfavorable feedback they solicited in the middle of the semester, the more their roommates derogated them and convinced them to make plans to find another roommate at the end of the semester. Furthermore, if people suspect that someone does not perceive them in a manner that befits their self-views, they will redouble their efforts to acquire self-verifying reactions. As noted earlier, in one study, participants who perceived themselves as either likable or dislikable learned that they would be interacting with someone who probably found them likable or dislikable. When participants suspected that their partner saw them either more or less favorably than they perceived themselves, they ramped-up their efforts to elicit self-verifying evaluations (e.g., Brooks et al., 2009; Swann and Hill, 1982; Swann and Read, 1981a, Study 2).

And what if people's efforts to obtain self-verifying evaluations fail? Even then, people may still cling to their self-views through the third strategy of self-verification: "seeing" nonexistent evidence. Self-views may guide at least three stages of information processing: attention, recall, and interpretation. For example, an investigation of selective attention revealed that participants with positive self-views spent longer scrutinizing evaluations they expected to be positive and people with negative self-views spent longer scrutinizing evaluations when they expected them to be negative (Swann and Read, 1981a, Study 1). Participants in a follow-up study displayed signs of selective recall. In particular, participants who perceived themselves positively remembered more positive than negative statements and those who perceived themselves negatively remembered more negative than positive statements. Finally, numerous investigations have shown that people tend to interpret information in ways [p. 28 ↓] that reinforce their self-views. Consider evidence that people with low self-esteem perceive their partners' sentiments toward them as being more negative than they actually are (e.g., Murray et al., 2000).

Together, attentional, encoding, retrieval, and interpretational processes may stabilize people's self-views by allowing them to "see" their worlds as offering more confirmation for their self-views than actually exists (for a review, see Swann et al., 2003c). These

strategies therefore represent a special case of the tendency for expectancies to channel information processing (e.g., Higgins and Bargh, 1987; Shrauger, 1975).

Interestingly, most investigations of self-verification processes have reported nearly symmetrical preferences of participants with positive and negative self-views. That is, just as participants with positive self-views displayed a preference for positive evaluations, participants with negative self-views displayed a preference for negative evaluations. In the early days of my research on self-verification, I had no idea how controversial this evidence would prove to be. I was soon to discover, however, that most of my colleagues were skeptical of the notion that people with negative self-views preferred negative evaluations. In fact, some of them would not buy a word of it.

The Backlash from Self-Enhancement Advocates

In the early 1980s, I noticed a baffling phenomenon. The more evidence for self-verification I published, the more skeptical my critics grew. The full magnitude of the problem, however, did not occur to me until an encounter with the great Stanley Schachter. After I had given a colloquium to the Psychology Department at Columbia (where he was the resident icon), I was excited to see him striding toward me. My excitement morphed into apprehension, however, when I noticed a scowl on his face. This was not just any scowl; it was so menacing that I instantly became convinced that he was about to take a swing at someone. Worse yet, judging from his trajectory, it seemed likely that that someone would be me. Stopping just short of my nose, he demanded, “So, are you telling me that people with negative self-concepts actually *want* negative evaluations?” I felt trapped. I sensed that if I caved, I would lose face, but if I stood my ground, I would lose my entire head. In the end I persuaded myself that I should hang tough, as my relatively youthful reflexes (he was more than twice my age) and wrestling experience would surely save me from serious injury. So convinced, I answered “At some level, yes” and prepared to duck. He stared at me in disbelief; I defiantly stared back. After what seemed like an eternity (spectators later told me the entire interaction was less than a minute), he announced loudly “I don’t believe it” and marched off in a huff.

For a host of reasons, Schachter's reaction was deeply troubling. It was bad enough that one of the world's most eminent social psychologists found my findings unpersuasive. More worrisome was the possibility that his concerns represented the tip of a much more ominous iceberg. Indeed, I would soon realize that for an increasingly vocal group of critics, my findings were not simply counterintuitive; they had been thoroughly discredited more than a decade earlier. The focal point of their concerns was an early study by Aronson and Carlsmith (1962). In this study, the experimenter asked a group of Harvard students to determine if the people pictured in series of photographs suffered from schizophrenia. After each of 100 trials, he delivered either positive or negative feedback to subjects. The crucial group received predominantly negative feedback for the first 80 trials followed by positive feedback on the last 20 trials. Shortly thereafter the experimenter indicated that there had been an oversight and asked subjects to take the final 20 trials of the test again.

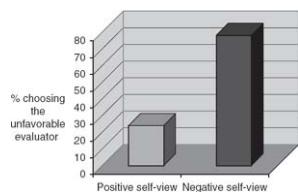
Aronson and Carlsmith's (1962) dependent measure was the extent to which subjects [p. 29 ↓] modified their responses to the final trials. Surprisingly, those who received unexpectedly positive feedback undermined their good fortune by modifying their responses! Theoretically, 80 trials of negative feedback had caused these participants to develop negative self-conceptions so that the positive feedback on the final trials produced dissonance. They accordingly altered their responses on the last 20 trials to reduce the dissonance created by the unexpectedly positive feedback.

Unfortunately, the results of the Aronson and Carlsmith study proved to be as difficult to replicate as they were provocative, with only 4 of 17 replication attempts succeeding (Dipboye, 1977). This rather dismal track record was enough to convince most people that Aronson and Carlsmith's findings were a fluke. More generally, critics argued that in a fair fight, self-consistency strivings were no match for self-enhancement strivings. This belief remains firmly entrenched among many social psychologists to this day, with most contemporary theorists tending either to subsume self-consistency strivings within a self-enhancement perspective (e.g., Schlenker, 1985; Sedikides and Gregg, 2008; Steele, 1988; Tesser, 1988) or to ignore them altogether.

My critics, noting a superficial similarity between the Aronson and Carlsmith (1962) findings and self-verification effects, dismissed evidence for self-verification. This was misguided, for it is inappropriate to link the two sets of findings. Most important, if

one looks closely at the procedures employed in the two sets of studies, one sees a crucial difference. In the self-verification studies, the experimenters *measured* the self-concepts of participants. This allowed them to tap into people's desire for self-stability and coherence. In contrast, Aronson and Carlsmith sought to *manipulate* self-views (by presenting participants with feedback indicating that they were unable to diagnose schizophrenics). Surely, providing negative feedback to a 20-year-old Harvard student is not likely to convince him that he does not know a crazy person when he sees one. For this reason, such a manipulation may put people in a bad mood, but it will not produce the chronic negative self-views needed to motivate self-verification strivings.

*Figure 27.1 Desire to interact with a negative evaluator as a function of self-view.
(Adapted from Swann et al. (1992))*



From this perspective, difficulties in replicating the Aronson and Carlsmith findings have no bearing on the replicability of self-verification effects. And, in fact, subsequent research bolstered this conclusion. Indeed, over the next several years, researchers in other labs and my own students replicated the basic self-verification effect (i.e., people with negative self-views preferred and sought negative over positive evaluations) dozens of times (e.g., Hixon and Swann, 1993; Robinson and Smith-Lovin, 1992; Swann et al, 1989, 1990, 1992c, 1992d). [Figure 27.1](#) shows an exemplary set of findings: just as [p. 30 ↓] people with positive self-views preferred to interact with a positive evaluator, people with negative self-views preferred to interact with someone who evaluated them negatively. Further, people with negative self-views seem to be truly drawn to self-verifying interaction partners rather than simply avoiding nonverifying ones. For example, when given the option of being in a different experiment, people with negative self-views chose to interact with a negative evaluator over participating in another experiment. Similarly, they chose being in a different experiment over interacting with a positive evaluator (Swann et al., 1992c).

Both men and women displayed this propensity, whether or not the self-views were easily changed and whether the self-views were associated with qualities that were specific (intelligence, sociability, dominance) or global (self-esteem, depression). People were particularly likely to seek self-verifying evaluations if their self-views were confidently held (e.g., Pelham and Swann, 1994; Swann and Ely, 1984; Swann et al., 1988) and important (Swann and Pelham, 2002), or extreme (Giesler et al., 1996). Moreover, in recent years researchers have shown that people also strive to verify negative (and positive) self-views associated with group membership. Such strivings emerge for both *collective* self-views (which are identities that characterize the person as well as the typical group member; Chen et al., 2004) and *group identities*, which refer to qualities of typical group members that may or may not characterize individual group members; Gómez et al., in press; Lemay and Ashmore, 2004).

In the face of such converging evidence, most adherents of the assumption that self-enhancement is the prepotent motivator of human behavior eventually relinquished their assertion that self-verification effects were not robust. Instead, they began to assert that the tendency for people with negative self-views to prefer and seek negative evaluations is counter-intuitive and bizarre. To counter such claims, I realized that I needed to show why people seek self-verification.

Why People Self-Verify

It is obvious why people work to maintain some negative self-views. After all, everyone possesses flaws and weaknesses and it makes perfect sense to develop and maintain negative self-views that correspond to these flaws and weaknesses. For example, people who lack some ability (as in those who are tone-deaf or cannot jump) will have numerous reasons for bringing others to recognize their shortcomings. For instance, when the appraisals of relationship partners square with objective reality, such partners will develop realistic expectations that the person can confirm and thus avoid disappointing the partner.

The adaptiveness of self-verification strivings, however, is much less obvious when people develop globally negative self-views (e.g., "I am worthless") that have no clear objective basis. Active efforts to maintain such negative self-views by, for example,

gravitating toward harsh or abusive partners, is surely maladaptive. At the very least, such activities seem to directly contradict the predictions of one of social psychology's most prominent approaches, self-enhancement theory. In fact, one of the greatest challenges to self-verification researchers is understanding how the motive interacts with the self-enhancement motive (e.g., Kwang and Swann, 2009).

Self-Enhancement versus Self-Verification

Self-enhancement theory can be traced back at least as far as Allport (1937). By positing a vital and universal human need to view oneself positively, Allport sowed the seeds for what would develop into a patchwork of loosely related propositions dubbed "self-enhancement theory" (Jones, 1973). Today this theory has received considerable support, including evidence that people are motivated to obtain, maintain, and increase positive self-regard. There are [p. 31 ↓] also indications that the desire for self-enhancement is truly fundamental. First, there is the apparent ubiquity of this desire. Whether one examines people's social judgments, attributions, or overt behaviors, there appears to be a widespread tendency for them to favor themselves over others (for a review, see Leary, 2007). Second, traces of a preference for positivity emerge at a tender age. Indeed, within mere weeks of developing the ability to discriminate facial characteristics, five-month-olds attend more to smiling faces than to nonsmiling ones (Shapiro et al., 1987). Similarly, as early as four-and-a-half months of age, children preferentially orient to voices that have the melodic contours of acceptance (Fernald, 1993). Third, among adults, a preference for positive evaluations emerges before other preferences (Swann et al., 1990). In particular, when forced to choose between two evaluators quickly, participants selected the positive evaluator even if they viewed themselves negatively. Only when given time to reflect did participants with negative self-views choose the negative, self-verifying partner.

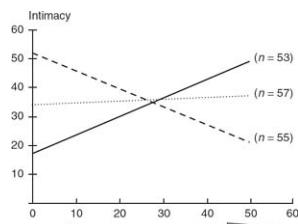
Yet, as potent as the desire for positivity may be, the results summarized earlier in this chapter indicate that self-verification strivings are quite robust. In fact, contrary to self-enhancement theory, people with negative self-views display a clear tendency to seek and embrace negative rather than positive partners. Furthermore, although the early demonstrations of self-verification strivings were conducted in the laboratory, later field studies showed a parallel pattern that was, in many respects, even more

remarkable than the initial studies. The first study in this series was designed to compare how people with positive self-views and negative self-views react to marital partners whose appraisals varied in positivity (Swann et al., 1994). The investigators recruited married couples who were either shopping at a local mall or horseback riding at a ranch in central Texas. The researchers approached potential participants and invited them to complete a series of questionnaires. They began with the Self-Attributes Questionnaire (SAQ; Pelham and Swann, 1989), a measure that focused on five attributes that most Americans regard as important: intelligence, social skills, physical attractiveness, athletic ability, and artistic ability. Then participants completed it again. This time, however, they rated their spouse. Finally, husbands and wives completed a measure of their commitment to the relationship. While each person completed these questionnaires, his or her spouse completed the same ones. The researchers thus had indices of what everyone thought of themselves, what their spouses thought of them, and how committed they were to the relationship.

How did people react to positive or negative evaluations from their spouses? As shown in [Figure 27.2](#), people with positive self-views responded in the intuitively-obvious way—the more favorable their spouses were, the more committed they were. By contrast, people with negative self-views displayed the opposite reaction; the more favorable their spouses were, the *less* committed they were. Those with moderate self-views were most committed to spouses who appraised them moderately.

Subsequent researchers attempted to replicate this effect (e.g., Cast and Burke, 2002; De La Ronde and Swann, 1998; Murray et al., 2000; Ritts and Stein, 1995; Schafer et al., 1996). Although the strength of the effect varied, each study reported some evidence that people preferred self-verifying spouses, even if their self-views were negative. A meta-analysis revealed that among married persons, the self-verification effect was stronger than the self-enhancement effect (Kwang and Swann, 2010). Moreover, a parallel finding emerged in a study of college student roommates (Swann and Pelham, 2002). Nevertheless, rather than accepting such findings as evidence of a desire for self-verification, advocates of self-enhancement theory refused to give up the fight. Instead, they insisted that what appeared to be self-verification strivings were, ironically, self-enhancement strivings gone awry.

Figure 27.2 Marital intimacy as a function of participants self-views and spouses' appraisals. (Based on Swann et al. (1994))



[p. 32 ↓]

Are Self-Verification Strivings Actually Self-Enhancement Strivings in Disguise?

One variation on this argument has been that self-verification effects are driven by a tiny segment of the population who suffer from flawed personalities such as masochism or self-destructive tendencies. From this vantage point, it was the personality flaw rather than the negative self-view that caused people with negative self-views to embrace negative evaluations and evaluators.

One counter to such claims is offered by an interesting aspect of the results of the investigation of married couples described above. Careful inspection of the findings revealed that it was not just persons with negative self-views who eschewed overly positive evaluations, for even people with positive self-views displayed less commitment to spouses whose evaluations were *extremely* favorable (Swann et al., 1994). Thus, the self-verification effect was not restricted to people with negative self-views; anyone who sensed that a spouse appraised them in an overly favorable manner tended to withdraw from the relationship.

Although these data are consistent with a self-verification explanation, they do not explicitly show that it was the *self-views* of people who thought poorly of themselves that caused them to choose negative evaluators. In search of such evidence, we

(Swann et al., 1990) hypothesized that there were differences in the cognitive operations that gave rise to self-enhancement versus self-verification strivings. In principle, self-enhancement strivings seem to require only one step: upon classifying the evaluation, people embrace positive evaluations and reject negative evaluations. In contrast, self-verification strivings logically require at least two steps. After classifying the evaluation, it needed to be compared to the self-view, for only then could the person choose to embrace verifying evaluations and avoid nonverifying ones. With this reasoning in hand, we predicted that depriving people of cognitive resources while they were choosing an interaction partner would interfere with their ability to access their self-concept. As a result, people who might ordinarily self-verify would self-enhance instead (cf. Paulhus and Levitt, 1987).

We tested these ideas by depriving participants of cognitive resources. In one study we did this by having people rehearse a phone number. While they struggled not to forget the phone number, they chose between a positive or negative evaluation. Deprived of the cognitive resources, they needed to compare [p. 33 ↓] the evaluation with their self-view; people with negative self-views suddenly behaved like their positive self-view compatriots – they chose positive evaluations over negative ones. When these same participants were later given several moments to access their self-views, however, they chose the negative, self-verifying evaluations. Later studies replicated this effect using other manipulations of resource deprivation, such as having participants choose partners hurriedly (Hixon and Swann, 1993). By showing that it was the ability to access their negative self-views that caused participants to choose negative evaluators, the resource deprivation studies showed that self-views rather than “flawed personalities” underlay self-verification strivings.

Another way of testing the flawed personalities hypothesis was to determine what people were thinking as they chose an interaction partner. To this end, we (Swann et al., 1992b) conducted a “think-aloud” study. People with positive and negative self-views thought out loud into a tape recorder as they chose an evaluator to interact with. As in the earlier studies, people with positive self-views tended to choose the positive evaluator and people with negative self-views tended to choose the negative evaluator. Of greatest relevance here, subsequent analyses of the tape recordings revealed no evidence that masochism or self-destructive tendencies drove the self-verifying choices of participants. To the contrary, participants with negative self-views seemed torn and

ambivalent as they chose negative partners. One person with negative self-views, for example, noted that:

I like the [favorable] evaluation but I am not sure that it is, ah, correct, maybe. It *sounds* good, but [the negative evaluator] ... seems to know more about me. So, I'll choose [the negative evaluator].

The think-aloud study also provided direct support for self-verification theory. The remarks of self-verifiers – both those with negative self-views who chose negative partners and those with positive self-views who chose favorable partners – indicated that they preferred partners who made them feel that they knew themselves. Consistent with self-verification theory, they were concerned with the match between the partner's evaluation and what they knew to be true of them:

Yeah, I think that's pretty close to the way I am. [The negative evaluator] better reflects my own view of myself, from experience.

There was also evidence that pragmatic considerations contributed to self-verification strivings, with self-verifiers voicing a concern with getting along with the evaluators during the forthcoming interaction:

Since [the negative evaluator] seems to know my position and how I feel sometimes, maybe I'll be able to get along with him.

In short, the results of the think-aloud study indicated that both epistemic and pragmatic considerations motivated participants to choose partners whose evaluations confirmed their self-views. As I will show below, the results of the think-aloud study, together with the marriage partner study, were also useful in addressing the possibility that people sought negative evaluations in a misguided effort to obtain positive evaluations.

Perceptiveness of the Evaluator

The distinction between desiring an evaluator who seems perceptive versus one who bolsters one's feelings of coherence parallels the difference between buying a car because it looks sporty versus choosing it because it makes one feel admired. In the

think-aloud study, people who mentioned a concern with perceptiveness focused on qualities of the evaluator, such as being “on the ball” or “insightful.” In contrast, people who emphasized coherence stressed a concern with feeling that the evaluator made them feel that they knew themselves. Those who mentioned being concerned with the perceptiveness of the evaluator were not the same ones who expressed coherence-related concerns, [p. 34 ↓] indicating that the two sets of concerns were independent. In addition, results of the marriage partner study indicated that relationship quality was driven by the extent to which the spouse was self-confirming rather than perceptive. In particular, commitment to relationships correlated with confidence that their spouses’ appraisals would make them “feel that they really knew themselves” rather than “confused them.” Commitment was not related with estimates of the perceptiveness of spouses, however.

Self-Improvement

Another rival explanation was that people with negative self-views choose interaction partners who thought poorly of them because they believed that such partners might give them critical feedback that would help them improve themselves. Participants in the think-aloud study did not mention this possibility, however. The results of the marital partners study also countered this possibility. When asked if they thought their spouse would provide them with information that would enable them to improve themselves, people with negative self-views were decidedly pessimistic, thus arguing against the possibility that this motive drew them into self-verifying relationships.

Perceived Similarity

Considerable evidence indicates that people prefer those who have similar values and beliefs. For example, people typically prefer their friends and associates who share their political beliefs, tastes in music, and the like (Byrne, 1971). Given this, it may be that people find self-verifying partners appealing because they suspect that such partners will agree with them on topics and issues that are unrelated to who they are. Contrary to this possibility, participants in the think-aloud study scarcely mentioned the partners’

likely attitudes. The results of the marital partner study also provided no evidence that people's affinity for self-verifying partners reflected an effort to align themselves with spouses possessing similar attitudes.

Winning Converts

Converting an enemy into a friend is generally difficult, so pulling off such a stunt ought to be especially gratifying. Conceivably, this is what people with negative self-views had on their minds when they chose partners who viewed them negatively. In fact, several participants in the think-aloud study did allude to a desire to win over a partner, as evidenced by comments such as, "I kind of think that [the negative evaluator] is ... the kind of guy or girl I'd like to meet and I would like to show them." Yet, it was only people with *positive* self-views who mentioned this concern; people with negative self-views never brought it up. This stands to reason, as people with negative self-views surely lack confidence that they can readily turn an enemy into a friend.

The marriage partner study provided further ammunition against the "winning converts" hypothesis. If people with negative self-views wished to "convert" a spouse who was initially critical, they should have expressed the most interest in partners whose evaluations of them seemed likely to grow more favorable over the course of the relationship. To the contrary, people with negative self-views tended to commit themselves more to spouses whose evaluations they expected to grow slightly more *negative* over time. Clearly, people with negative self-views choose rejecting interaction partners for very different reasons than people with positive self-views did.

Self-Verification versus Accuracy

Some critics have asserted that evidence of self-verification processes is unsurprising because people with negative self-views are merely seeking evaluations that confirm actual deficiencies. Let me begin by acknowledging that people with negative self-views undoubtedly possess *some* negative qualities. Tragically, people sometimes develop the conviction that they are flawed when in reality they are not. Support for this idea comes from research in which the researchers examined the feedback-

seeking activities of [p. 35 ↓] people who were clinically depressed (Giesler et al., 1996). Depressed people regarded negative evaluations to be especially accurate and were more apt to seek them. This finding is significant because there is no evidence that depressed people actually possess chronic deficiencies that would justify their quest for negative feedback. Similarly, it is difficult to imagine a convincing justification for the tendency for people with low self-esteem to feel that they are worthless and undeserving of love. Finally, if depressed persons were truly as deficient as their negative self-views would suggest, one would expect that their negative self-views would remain this way on a more-or-less permanent basis. They do not: once depression clears, the self-views of formerly depressed people bounce back to normal.

Note that I am not suggesting that people have no interest in winning the approval of their relationship partners. Indeed, the self-verification process requires that relationships survive, for there can be no self-verification if there is no relationship. For this reason, people are highly motivated to bring their relationship partners to see them positively on qualities that are essential to the survival of the relationship. Physical attractiveness is one such quality. Not surprisingly, target persons not only want their dating partners to see them as much more attractive than they see themselves, they actually take steps to ensure their partners view them this way (e.g., Swann et al., 2002). Moreover, such steps are effective, for people's partners actually develop appraisals that verify targets' more-attractive-than-usual selves. Apparently, people with negative self-views recognize that for their relationships to remain viable, they must be perceived in a relatively positive manner on relationship-relevant dimensions. We dubbed this phenomenon "strategic self-verification," as people gained verification for strategic selves that differed from their chronic selves.

How can evidence of strategic self-verification be reconciled with the research discussed earlier indicating that people seek and elicit self-verifying evaluations? Apparently, people with negative self-views prefer and seek negative evaluations regarding characteristics that are low in relationship-relevance (e.g., intelligence, artistic), presumably because verification of such negative qualities will not threaten the viability of the relationship. At the same time, on dimensions that are critical to the relationship, they strive to acquire evaluations that are more positive than those that they typically receive but which verify the self that they have presented to their partners. In this way, targets may receive verification of qualities that are low in relationship

relevance as well as verification of circumscribed, highly positive selves that they negotiate with their partner on qualities that are high in relationship relevance (cf. Neff and Karney, 2005).

Interestingly, this evidence for the moderating role of relationship-relevance is consistent with self-verification theory's notion that people strive for convergence between their self-views and the social realities that maintain them. Nevertheless, it is inconsistent with the theory's assumption that people strive to negotiate identities that match their chronic self-views (Swann, 1983). Apparently, people will seek verification of their negative self-views only if doing so does not risk being abandoned, for abandonment would completely cut off the supply of verification (cf. Hardin and Higgins's, 1996, discussion of people's unwillingness to embrace epistemic truth if it undermines the relationship aspect of shared realities). While enacting such relationship-specific selves departs from the assumptions of classical trait and self theory, it is quite consistent with Mischel and Shoda's (1999) notion that people strive for intra-individual consistency and with my suggestion that people strive for circumscribed accuracy (e.g., Gill and Swann, 2004; Swann, 1984). It is also consistent with conceptions of the self in East Asia in which people eschew self-descriptions that emphasize abstract traits in favor of self-views that emphasize responsiveness to social roles [p. 36 ↓] and cross-situational flexibility (e.g., Choi and Choi, 2002; Kanagawa et al., 2001; for a discussion, see English et al., 2008).

Returning to the more general point here, our research has uncovered little support for various ironic explanations of self-verification strivings. Instead, it appears that a desire for self-stability and associated feelings of coherence motivates people to strive for self-verification. If self-verification strivings are indeed built into our psychological architecture, one would expect two things. First, self-verification strivings should act as a powerful counterpoint to self-enhancement strivings. A recent meta-analysis supports this possibility, indicating that self-verification strivings trumped self-enhancement strivings on measures of feedback seeking and relationship quality while self-enhancement strivings prevailed only when researchers focused on affective reactions (Kwang and Swann, 2010). Second, researchers should find that self-verification is associated with various personal and social benefits.

The Personal and Social Psychological Utility of Self-Verification

There is growing evidence that self-verification strivings predict a host of important outcomes. These outcomes occur at several different levels of analysis, including the individual, interpersonal, and societal level of analysis.

Individual Outcomes

For the roughly 70 percent of people who have positive self-views (e.g., Diener and Diener, 1995), the case for the personal adaptiveness of self-verification strivings is clear and compelling. Self-verification strivings bring stability to people's lives, rendering their experiences more coherent, orderly, and comprehensible than they would be otherwise. Success in acquiring self-verifying evaluations may bring with it important psychological benefits. For example, insofar as people's partners are self-verifying, their relationships will be more predictable and manageable. Such predictability and manageability may not only enable people to achieve their relationship goals (e.g., raising children, coordinating careers), it may also be psychologically comforting and anxiety reducing.

For people with negative self-views, however, the fruits of self-verification strivings are adaptive in some instances but not in others. In most instances, seeking verification for negative self-views will be adaptive when such views accurately reflect immutable personal limitations (e.g., lack of height). Despite contentions to the contrary (Taylor and Brown, 1988), there is no convincing evidence that self-delusions are adaptive (Kwang and Swann, 2010).

The picture is much cloudier, however, when people develop *inappropriately* negative self-views – that is, self-views that exaggerate or misrepresent their limitations (e.g., believing that one is fat when one is thin, or dull witted when one is bright). On the positive side, eliciting negative but self-verifying evaluations has the virtue of holding anxiety at bay. For example, one set of investigators (Wood et al., 2005) contrasted

the reactions of high and low self-esteem participants to success experiences. Whereas high self-esteem persons reacted quite favorably to success, low self-esteem participants reported being anxious and concerned, apparently because they found success to be surprising and unsettling (cf. Lundgren and Schwab, 1977). Similarly, others (Ayduk et al., 2008) observed participants' cardiovascular responses to positive and negative evaluations. When people with negative self-views received positive feedback, they were physiologically "threatened" (distressed and avoidant). In contrast, when they received negative feedback, participants with negative [p. 37 ↓] self-views were physiologically "challenged" or "galvanized" (i.e., cardiovascularly aroused but in a manner associated with approach motivation). The opposite pattern emerged for people with positive self-views.

If people with negative self-views are stressed by positive information, over an extended period such information might actually produce debilitation. Empirical support for this possibility comes from several independent investigations. An initial pair of prospective studies (Brown and McGill, 1989) compared the impact of positive life events on the health outcomes of people with low versus high self-esteem. Positive life events (e.g., improvement in living conditions, getting good grades) predicted increases in health among high self-esteem participants but decreases in health among people low in self-esteem. Recent investigations by Shimizu and Pelham (2004) replicated and extended these results while controlling for negative affectivity, thereby undercutting the rival hypothesis that negative affect influenced both self-reported health and reports of symptoms. Remarkably, in all of these studies, positive life events were apparently so unsettling to people with low self-esteem that their physical health suffered.

But if receiving verification for negative self-views may be beneficial in some respects, the costs may outweigh the benefits in cases in which the self-views are more negative than warranted by objective reality. For instance, self-verification strivings may prompt people with negative self-views to gravitate toward partners who mistreat them, undermine their feelings of self-worth, or even abuse them. Once ensconced in such relationships, people may be unable to benefit from therapy because returning home to a self-verifying partner may undo the progress that was made in the therapist's office (Swann and Predmore, 1985). And the workplace may offer little solace, for the feelings of worthlessness that plague people with low self-esteem may make them ambivalent about receiving fair treatment, ambivalence that may undercut their

propensity to insist that they get what they deserve from their employers (Wiesenfeld et al., 2007). Moreover, such tragic outcomes are not limited to global negative self-views. As mentioned above, people who are thin sometimes develop the mistaken impression that they are fat, a perception that gives rise to *anorexia*, a major killer of teenage girls (Hoek, 2006). Clearly, for those who develop erroneous negative self-views, it is important to take steps to disrupt the self-verifying cycles in which they are often trapped (Swann, 1996; Swann et al., 2006). More generally, such instances illustrate how the process of self-verification can sometimes have negative consequences even though it is adaptive for most people most of the time.

Interpersonal Outcomes

Earlier, I speculated that during human evolutionary history, self-verification strivings may have increased inclusive fitness by making successful self-verifiers more predictable to other group members. Modern humans may benefit from self-verification strivings for similar reasons. In fact, research indicates that when members of small groups receive self-verification from other group members, their commitment to the group increases and performance improves (Swann et al., 2000, 2004).

Self-verification processes seem to be especially useful in small groups composed of people from diverse backgrounds. That is, out of a fear that they will be misunderstood, members of diverse groups may often refrain from expressing controversial ideas. Self-verification may reduce such fear by convincing them that they are understood. For this reason, they may open up to their coworkers. Such openness may, in turn, lead them to express off beat ideas that lead to creative solutions to problems. Performance may benefit (Polzer et al., 2002; Seyle et al., 2009).

[p. 38 ↓]

Societal Outcomes

Self-verification processes are also adaptive for groups and the larger society. Because self-verification processes make people predictable to one another, they may grease

the wheels of social interaction. Self-verification processes seem to be especially useful in small groups composed of people from diverse backgrounds. In fact, when group members offer one another self-verification, relatively diverse groups actually outperform relatively nondiverse groups – an instance in which the “value in diversity hypothesis” seems to hold true (e.g., Polzer et al., 2002; Swann et al., 2004).

Self-verification can also help eradicate social stereotypes. In small groups, those who offer other group members self-verification are more apt to individuate them – that is, recognize them as unique individuals rather than as exemplars of social stereotypes (Swann et al., 2003a). Over time, such treatment could influence targets and perceivers alike. Targets who are treated as unique individuals will be encouraged to develop qualities that reflect their idiosyncratic competences and capacities. At the same time, perceivers who individuate other group members will relinquish their social stereotypes (Swann et al., 2003b).

There is also evidence that self-verification strivings may play a role in extreme behaviors. In a recent series of studies, investigators identified a group of people whose personal identities were “fused” with their social identities (Swann et al., 2009). Because the personal and social self are functionally equivalent among such individuals, activating one is tantamount to activating the other. Consistent with this, when we activated a personal self by challenging its validity, people displayed compensatory self-verification strivings. Among fused persons, such compensatory activity took the form of increased willingness to perform extraordinary behaviors, such as dying for the group (see also, Gómez et al., *in press*; Swann, et al., 2010a, 2010b).

New Directions

Current research on self-verification is moving in several distinct directions. One approach focuses on tradeoffs between self-verification and other motives such as positivity, particularly in close relationships (e.g., Neff and Karney, 2005). One fascinating issue here is how people create and sustain idiosyncratic social worlds that are disjunctive with the worlds that they have created outside the relationship (Swann et al., 2002). Another emerging theme (e.g., Chen et al., 2004; Gómez et al., 2009) has been on the verification of social identities (i.e., identities associated

with the groups people align themselves with, such as Democrat, American, etc.) as compared with personal identities (i.e., self-views referring to personal qualities, such as intelligent, athletic, etc.). A third set of questions have emerged regarding similarities and differences in the way that self-verification strivings unfold in other cultures (English et al., 2008). My take on this issue is that all people desire coherence and predictability but that this desire may express itself differently depending upon the extent to which the culture values selves that are cross-situationally consistent (e.g., Western culture) or relationship specific (e.g., some Asian cultures).

Much of my own recent work has focused on the interplay of self-verification strivings and identity negotiation, the processes whereby people in relationships reach agreements regarding “who is who.” Identity negotiation theory (Swann and Bosson, 2008) integrates self-verification theory's emphasis on the activities of targets of social perception with behavioral confirmation theory's (Snyder and Swann, 1978b) emphasis on the activities of perceivers. My recent interest in identity negotiation theory has brought me full circle, as I am once again examining the impact of interpersonal expectancies, as I did as a graduate student. This time around, however, I can exploit the knowledge gained during three decades of research on [p. 39 ↓] self-verification processes. At the very least, I feel that I now have some insight into the nature and consequences of the negative identities that Tommy negotiated with his peers and the staff at Camp Sunshine.

Acknowledgments

I thank Rebecca Caldwell and Tory Higgins for helpful comments on an earlier draft of the article.

References

- Allport, G.W. (1937) Personality: A Psychological Interpretation . New York: Holt.

Aronson, E. (1968) A theory of cognitive dissonance: A current perspective . In L. Berkowitz (ed.), *Advances in Experimental Social Psychology* , 4, 1–34. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960075-1>

Aronson, E. and Carlsmith, J.M. Performance expectancy as a determinant of actual performance . *Journal of Abnormal and Social Psychology* , 65, 178–182. (1962) <http://dx.doi.org/10.1037/h0042291>

Ayduk, O., Gyurak, A., Akinola, M. and Mendes, W.B. (2011) Self-verification processes revealed in implicit and behavioral responses to feedback . Manuscript under review, UC Berkeley.

Brooks, M.L., Mehta, P.H. and Swann, W.B., Jr. (2009) Reclaiming the self: Compensatory self-verification following a deprivation experience . Unpublished manuscript, University of Texas at Austin.

Brown, J.D. and McGill, K.J. The cost of good fortune: When positive life events produce negative health consequences . *Journal of Personality and Social Psychology* , 55, 1103–1110. (1989) <http://dx.doi.org/10.1037/0022-3514.57.6.1103>

Byrne, D. (1971) *The Attraction Paradigm* . New York: Academic Press.

Cassidy, J., Ziv, Y., Mehta, T.G. and Feeney, B.C. Feedback seeking in children and adolescents: associations with self-perceptions, representations, and depression . *Child Development* , 74, 612–628. (2003) <http://dx.doi.org/10.1111/1467-8624.7402019>

Cast, A.D. and Burke, P.J. A theory of self-esteem . *Social Forces* , 80, 1041–1068. (2002) <http://dx.doi.org/10.1353/sf.2002.0003>

Chang-Schneider, C. and Swann, W.B., Jr. (2009) Wearing self-esteem like a flag: Conveying our high- and low-self-esteem to others . Unpublished manuscript, University of Texas at Austin.

Chen, S., Chen, K.Y. and Shaw, L. Self-verification motives at the collective level of self-definition . *Journal of Personality and Social Psychology* , 86, 77–94. (2004) <http://dx.doi.org/10.1037/0022-3514.86.1.77>

Choi, I. and Choi, Y. Culture and self-concept flexibility . Personality and Social Psychology Bulletin , 28, 1508–1517. (2002) <http://dx.doi.org/10.1177/014616702237578>

Cooley, C.S. (1902) Human Nature and the Social Order . New York: Scribner's.

De La Ronde, C. and Swann, W.B., Jr. Partner verification: Restoring shattered images of our intimates . Journal of Personality and Social Psychology , 75, 374–382. (1998) <http://dx.doi.org/10.1037/0022-3514.75.2.374>

Diener, E. and Diener, M. Cross-cultural correlates of life satisfaction and self-esteem . Journal of Personality and Social Psychology , 68, 653–663. (1995) <http://dx.doi.org/10.1037/0022-3514.68.4.653>

Dipboye, R.L. A critical review of Korman's self-consistency theory of work motivation and occupational choice . Organizational Behavior and Human Performance , 18, 108–126. (1977) <http://dx.doi.org/10.1016/0030-5073%2877%2990021-6>

English, T., Chen, S. and Swann, W.B., Jr. (2008) A cross-cultural analysis of self-verification motives . In R. Sorrentino, ed. and S. Yamaguchi (eds), Handbook of Motivation and Cognition Across Cultures , pp. 119–142, San Diego: Elsevier. <http://dx.doi.org/10.1016/B978-0-12-373694-9.00006-4>

Fernald, A. Approval and disapproval: Infant responsiveness to verbal affect in familiar and unfamiliar languages . Child Development , 64, 657–674. (1993) <http://dx.doi.org/10.2307/1131209>

Festinger, L. (1957) A Theory of Cognitive Dissonance , Evanston: Row, Peterson.

Giesler, R.B., Josephs, R.A. and Swann, W.B., Jr. Self-verification in clinical depression: The desire for negative evaluation . Journal of Abnormal Psychology , 105, 358–368. (1996) <http://dx.doi.org/10.1037/0021-843X.105.3.358>

Gill, M.J. and Swann, W.B., Jr. On what it means to know someone: A matter of pragmatics . Journal of Personality and Social Psychology , 86, 405–418. (2004) <http://dx.doi.org/10.1037/0022-3514.86.3.405>

Goffman, E. (1959) *The Presentation of Self in Everyday Life* . Garden City, NY: Doubleday-Anchor.

Gómez, Á., Brooks, M.L., Buhrmester, M.D., Vázquez, A., Jetten, J. and Swann, W.B., Jr. (in press) On the nature of identity fusion: Insights into the construct and a new measure . *Journal of Personality and Social Psychology* .

Gómez, Á., Seyle, C., Huici, C. and Swann, W.B., Jr. Can self-verification strivings fully transcend the self-other barrier? Seeking verification of ingroup identities . *Journal of Personality and Social Psychology* , 97, 1021–1044. (2009)

Gosling, S. (2008) *Snoop: What Your Stuff Says About You* . New York: Basic.

Hardin, C.D. and Higgins, E.T. (1996) Shared reality: How social verification makes the subjective objective . In E.T. Higgins, ed. and R.M. Sorrentino (eds), *Handbook of Motivation and Cognition: The Interpersonal Context* (Vol. 3) . New York: Guilford Press.

Higgins, E.T. and Bargh, J.A. (1987) Social cognition and social perception . In M.R. Rosenzweig, ed. and L.W. Porter (eds), *Annual Review of Psychology* , 38, 369–425. Palo Alto: Annual Reviews. <http://dx.doi.org/10.1146/annurev.ps.38.020187.002101>

Hixon, J.G. and Swann, W.B., Jr. When does introspection bear fruit? Self-reflection, self-insight, and interpersonal choices . *Journal of Personality and Social Psychology* , 64, 35–43. (1993) <http://dx.doi.org/10.1037/0022-3514.64.1.35>

Hoek, H.W. Incidence, prevalence and mortality of anorexia nervosa and other eating disorders . *Current Opinion Psychiatry* , 19, 389–94. (2006) <http://dx.doi.org/10.1097/YCO.0000228759.95237.78>

Jones, S.C. Self and interpersonal evaluations: Esteem theories versus consistency theories . *Psychological Bulletin* , 79, 185–199. (1973) <http://dx.doi.org/10.1037/h0033957>

Kanagawa, C., Cross, S. and Markus, H. 'Who am I?' The cultural psychology of the conceptual self . *Personality and Social Psychology Bulletin* , 27, 90–103. (2001) <http://dx.doi.org/10.1177/0146167201271008>

Kwang, T. and Swann, W.B., Jr. Do people embrace praise even when they feel unworthy? A review of critical tests of self-enhancement versus self-verification . *Personality and Social Psychology Review* , 14, 263–280. (2010) <http://dx.doi.org/10.1177/1088868310365876>

Leary, M.R. Motivational and emotional aspects of the self . Annual Review of Psychology , 58, 317–344. (2007) <http://dx.doi.org/10.1146/annurev.psych.58.110405.085658>

Leary, M.R. and Baumeister, R.F. (2000) The nature and function of self-esteem: Sociometer theory . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 32, 2–51. San Diego: Academic Press <http://dx.doi.org/10.1016/S0065-2601%2800%2980003-9>

Lecky, P. (1945) *Self-consistency: A Theory of Personality* . New York: Island Press.

Lemay, E.P. and Ashmore, R.D. (2004) Reactions to perceived categorization by others during the transition to college: Internalizaton of self-verification processes . *Group Processes and Interpersonal Relations* , 173–187. <http://dx.doi.org/10.1177/1368430204043722>

LundgrenD.C. and SchwabM.R. Perceived appraisals by others, self-esteem, and anxiety . *Journal of Psychology* , 97, 205–213. (1977) <http://dx.doi.org/10.1080/00223980.1977.9923964>

McCall, G.J. and Simmons, J.L. (1966) *Identities and Interactions: An Examination of Human Associations in Everyday Life* . New York: Free Press.

Mead, G.H. (1934) *Mind, Self and Society* . Chicago: University of Chicago Press.

Mischel, W. and Shoda, Y. (1999) Integrating dispositions and processing dynamics within a unified theory of personality: The Cognitive Affective Personality System (CAPS) . In L.A. Pervin, ed. and O. John (eds), *Handbook of Personality: Theory and Research* , 2, 197–218. New York: Guilford Press.

Murray, S.L., Holmes, J.G., Dolderman, D. and Griffin, D.W. What the motivated mind sees: Comparing friends' perspectives to married partners' views of each other . *Journal of Experimental Social Psychology* , 36, 600–620. (2000) <http://dx.doi.org/10.1006/jesp.1999.1417>

Neff, L.A. and Karney, B.R. To know you is to love you: The implications of global adoration and specific accuracy for marital relationships . *Journal of Personality and Social Psychology* , 88, 480–497. (2005) <http://dx.doi.org/10.1037/0022-3514.88.3.480>

Paulhus, D.L. and Levitt, K. Desirable responding triggered by affect: Automatic egotism? *Journal of Personality and Social Psychology* , 52, 245–259. (1987) <http://dx.doi.org/10.1037/0022-3514.52.2.245>

Pelham, B.W. and Swann, W.B., Jr. From self-conceptions to self-worth: The sources and structure of self-esteem . *Journal of Personality and Social Psychology* , 57, 672–680. (1989) <http://dx.doi.org/10.1037/0022-3514.57.4.672>

Pelham, B.W. and Swann, W.B., Jr. The juncture of intrapersonal and interpersonal knowledge: Self-certainty and interpersonal congruence . *Personality and Social Psychology Bulletin* , 20, 349–357. (1994) <http://dx.doi.org/10.1177/0146167294204002>

Polzer, J.T., Milton, L.P. and Swann, W.B., Jr. Capitalizing on diversity: interpersonal congruence in small work groups . *Administrative Science Quarterly* , 47, 296–324. (2002) <http://dx.doi.org/10.2307/3094807>

Pratt, M.G. and Rafaeli, A. Organizational dress as a symbol of multilayered social identities . *Academy of Management Journal* , 40, 862–898. (1997) <http://dx.doi.org/10.2307/256951>

Ritts, V. and Stein, J.R. Verification and commitment in marital relationships: An exploration of self-verification theory in community college students . *Psychological Reports* , 76, 383–386. (1995) <http://dx.doi.org/10.2466/pr0.1995.76.2.383>

Robinson, D.T. and Smith-Lovin, L. Selective interaction as a strategy for identity maintenance: An affect control model . *Social Psychology Quarterly* , 55, 12–28. (1992) <http://dx.doi.org/10.2307/2786683>

Sacks, O. (1985) *The Man Who Mistook His Wife for a Hat and Other Clinical Tales*. New York: Simon & Shuster.

Schafer, R.B., Wickrama, K.A.S. and Keith, P.M. Self-concept disconfirmation, psychological distress, and marital happiness. *Journal of Marriage and the Family*, 58, 167–177. (1996) <http://dx.doi.org/10.2307/353385>

Schlenker, B.R. (1985) Identity and self-identification. In B.R. Schlenker (ed.), *The Self and Social Life*, pp. 65–99. New York: McGraw-Hill.

Secord, P.F. and Backman, C.W. (1965) An interpersonal approach to personality. In B. Maher (ed.), *Progress in Experimental Personality Research*, 2, 91–125. New York: Academic Press.

Sedikides, C. and Gregg, A.P. Self-enhancement: Food for thought. *Perspectives on Social Psychology*, 3, 102–116. (2008) <http://dx.doi.org/10.1111/j.1745-6916.2008.00068.x>

Seyle, D.C., Athle, D. and Swann, W.B., Jr. (2009) Value in diversity and the self: Verifying self-views promotes group connectedness and performance in diverse groups. Unpublished manuscript, University of Texas at Austin.

Shapiro, B., Eppler, M., Haith, M. and Reis, H. (1987) An event analysis of facial attractiveness and expressiveness. Paper presented at the Society for Research in Child Development, Baltimore, MD.

Shimizu, M. and Pelham, B.W. The unconscious cost of good fortune: implicit and positive life events, and health. *Health Psychology*, 23, 101–105 (2004) <http://dx.doi.org/10.1037/0278-6133.23.1.101>

Shrauger, J.S. Responses to evaluation as a function of initial self-perceptions. *Psychological Bulletin*, 82, 581–596. (1975) <http://dx.doi.org/10.1037/h0076791>

Snyder, M. and Swann, W.B., Jr. Hypothesis testing processes in social interaction. *Journal of Personality and Social Psychology*, 36, 1202–1212. (1978a) <http://dx.doi.org/10.1037/0022-3514.36.11.1202>

Snyder, M. and Swann, W.B., Jr. Behavioral confirmation in social interaction: From social perception to social reality . *Journal of Experimental Social Psychology* , 14, 148–162. (1978b) <http://dx.doi.org/10.1016/0022-1031%2878%2990021-5>

Steele, C.M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self . In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, Vol. 21 , pp. 261–302. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960229-4>

Swann, W.B., Jr. (1983) Self-verification: Bringing social reality into harmony with the self . In J. Suls, ed. and A.G. Greenwald (eds), *Psychological Perspectives on the Self* , 2, 33–66, Hillsdale, NJ: Erlbaum.

Swann, W.B., Jr. The quest for accuracy in person perception: A matter of pragmatics . *Psychological Review* , 91, 457–477. (1984) <http://dx.doi.org/10.1037/0033-295X.91.4.457>

Swann, W.B., Jr. (1996) Self-traps: The Elusive Quest for Higher Self-esteem . Freeman: New York.

Swann, W.B., Jr. and Bosson, J. (2008) Identity negotiation: A theory of self and social interaction . In O. John, ed. , R. Robins, ed. , and L. Pervin (eds), *Handbook of Personality Psychology: Theory and Research I* , pp. 448–471. New York: Guilford Press.

Swann, W.B., Jr., Bosson, J.K. and Pelham, B.W. Different partners, different selves: The verification of circumscribed identities . *Personality and Social Psychology Bulletin* , 28, 1215–1228. (2002) <http://dx.doi.org/10.1177/01461672022812007>

Swann, W.B., Jr., Chang-Schneider, C.S. and McClarty, K.L. Do people's self-views matter? Self-concept and self-esteem in everyday life . *American Psychologist* , 62, 84–94. (2006) <http://dx.doi.org/10.1037/0003-066X.62.2.84>

Swann, W.B., Jr., De La Ronde, C. and Hixon, J.G. Authenticity and positivity strivings in marriage and courtship . *Journal of Personality and Social Psychology* , 66, 857–869. (1994) <http://dx.doi.org/10.1037/0022-3514.66.5.857>

Swann, W.B., Jr. and Ely, R.J. A battle of wills: Self-verification versus behavioral confirmation . *Journal of Personality and Social Psychology* , 46, 1287–1302. (1984) <http://dx.doi.org/10.1037/0022-3514.46.6.1287>

Swann, W.B., Jr., Gómez, Á., Seyle, C. and Morales, F. Identity fusion: The interplay of personal and social identities in extreme group behavior . *Journal of Personality and Social Psychology* , 96, 995–1011. (2009) <http://dx.doi.org/10.1037/a0013668>

Swann, W.B., Jr., Gómez, Á., Dovidio, J., Hart, S. and Jetten, J. Dying and killing for one's group: Identity fusion moderates responses to intergroup versions of the trolley problem . *Psychological Science* , 21, 1176–1183 (2010a) <http://dx.doi.org/10.1177/0956797610376656>

Swann, W.B., Jr., Gómez, Á., Huici, C., Morales, F. and Hixon, J.G. Identity fusion and self-sacrifice: Arousal as catalyst of pro-group fighting, dying and helping behavior . *Journal of Personality and Social Psychology* , 99, 824–841. (2010b) <http://dx.doi.org/10.1037/a0020014>

Swann, W.B., Jr. and Hill, C.A. When our identities are mistaken: Reaffirming self-conceptions through social interaction . *Journal of Personality and Social Psychology* , 43, 59–66. (1982) <http://dx.doi.org/10.1037/0022-3514.43.1.59>

Swann, W.B., Jr., Hixon, J.G. and De La Ronde, C. Embracing the bitter ‘truth’: Negative self-concepts and marital commitment . *Psychological Science* , 3, 118–121. (1992a) <http://dx.doi.org/10.1037/0022-3514.49.6.1609>

Swann, W.B., Jr., Hixon, J.G., Stein-Seroussi, A. and Gilbert, D.T. The fleeting gleam of praise: Behavioral reactions to self-relevant feedback . *Journal of Personality and Social Psychology* , 59, 17–26. (1990) <http://dx.doi.org/10.1037/0022-3514.59.1.17>

Swann, W.B., Jr., Kwan, V.S.Y., Polzer, J.T. and Milton, L.P. Vanquishing stereotypic perceptions via individuation and self-verification: Waning of gender expectations in small groups . *Social Cognition* , 21, 194–212. (2003a) <http://dx.doi.org/10.1521/soco.21.3.194.25341>

Swann, W.B., Jr., Kwan, V.S.Y., Polzer, J.T. and Milton, L.P. Fostering group identification and creativity in diverse groups: The role of individuation and self-verification . *Personality and Social Psychology Bulletin* , 29, 1396–1406. (2003b) <http://dx.doi.org/10.1177/0146167203256868>

Swann, W.B., Jr., Milton, L.P. and Polzer, J.T. Should we create a niche or fall in line? Identity negotiation and small group effectiveness . *Journal of Personality and Social Psychology* , 79, 238–250. (2000) <http://dx.doi.org/10.1037/0022-3514.79.2.238>

Swann, W.B., Jr. and Pelham, B.W. Who wants out when the going gets good? Psychological investment and preference for self-verifying college roommates . *Journal of Self and Identity* , 1, 219–233. (2002) <http://dx.doi.org/10.1080/152988602760124856>

Swann, W.B., Jr., Pelham, B.W. and Chidester, T. Change through paradox: Using self-verification to alter beliefs . *Journal of Personality and Social Psychology* , 54, 268–273. (1988) <http://dx.doi.org/10.1037/0022-3514.54.2.268>

Swann, W.B., Jr., Pelham, B.W. and Krull, D.S. Agreeable fancy or disagreeable truth? Reconciling self-enhancement and self-verification . *Journal of Personality and Social Psychology* , 57, 782–791. (1989) <http://dx.doi.org/10.1037/0022-3514.57.5.782>

Swann, W.B.Jr., Polzer, J.T., Seyle, C. and Ko, S. Finding value in diversity: Verification of personal and social self-views in diverse groups . *Academy of Management Review* , 29, 9–27. (2004)

Swann, W.B., Jr. and Predmore, S.C. Intimates as agents of social support: Sources of consolation or despair? *Journal of Personality and Social Psychology* , 49, 1609–1617. (1985) <http://dx.doi.org/10.1037/0022-3514.49.6.1609>

Swann, W.B., Jr. and Read, S.J. Self-verification processes: How we sustain our self-conceptions . *Journal of Experimental Social Psychology* , 17, 351–372. (1981a) <http://dx.doi.org/10.1016/0022-1031%2881%2990043-3>

Swann, W.B., Jr. and Read, S.J. Acquiring self-knowledge: The search for feedback that fits . *Journal of Personality and Social Psychology* , 41, 1119–1128. (1981b) <http://dx.doi.org/10.1037/0022-3514.41.6.1119>

Swann, W.B., Jr., Rentfrow, P.J. and Guinn, J. (2003c) Self-verification: The search for coherence . In M.Leary and J.Tagney, *Handbook of Self and Identity* , pp. 367–383. New York: Guilford Press.

Swann, W.B., Jr. and Snyder, M. On translating beliefs into action: Theories of ability and their application in an instructional setting . *Journal of Personality and Social Psychology* , 38, 879–888. (1980) <http://dx.doi.org/10.1037/0022-3514.38.6.879>

Swann, W.B., Jr., Stein-Seroussi, A. and Giesler, B. Why people self-verify . *Journal of Personality and Social Psychology* , 62, 392–401. (1992b) <http://dx.doi.org/10.1037/0022-3514.62.3.392>

Swann, W.B., Jr., Wenzlaff, R.M. and Tafarodi, R.W. Depression and the search for negative evaluations: More evidence of the role of self-verification strivings . *Journal of Abnormal Psychology* , 101, 314–371. (1992c) <http://dx.doi.org/10.1037/0021-843X.101.2.314>

Swann, W.B., Jr., Wenzlaff, R.M., Krull, D.S. and Pelham, B.W. The allure of negative feedback: Self-verification strivings among depressed persons . *Journal of Abnormal Psychology* , 101, 293–306. (1992d) <http://dx.doi.org/10.1037/0021-843X.101.2.293>

Taylor, S.E. and Brown, J.D. Illusion and well-being: A social psychological perspective on mental health . *Psychological Bulletin* , 103, 193–210. (1988) <http://dx.doi.org/10.1037/0033-2909.103.2.193>

Tesser, A. (1988) Toward a self-evaluation maintenance model of social behavior . In L. Berkowitz (ed.), *Advances in Experimental Psychology*, Vol. 21: Social Psychological Studies of the Self: Perspectives and Programs , pp. 181–227. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960227-0>

Wiesenfeld, B.M., Swann, W.B., Jr, Brockner, J. and Bartel, C. Is more fairness always preferred? Self-esteem moderates reactions to procedural justice . *Academy of Management Journal* , 50, 1235–1253. (2007) <http://dx.doi.org/10.2307/20159922>

Winkielman, P., Schwarz, N. and Nowak, A. (2002) Affect and processing dynamics: Perceptual fluency enhances evaluations . In S. Moore, ed. and M. Oaksford (eds),

Emotional Cognition: From Brain to Behaviour , pp. 111–136. Amsterdam: John Benjamins.

Wood, J.V., Heimpel, S.A., Newby-Clark, I. and Ross, M. Snatching defeat from the jaws of victory: Self-esteem differences in the experience and anticipation of success . Journal of Personality and Social Psychology , 89, 764–780. (2005) <http://dx.doi.org/10.1037/0022-3514.89.5.764>

<http://dx.doi.org/10.4135/9781446249222.n27>

Handbook of Theories of Social Psychology

Implicit Theories

Contributors: Carol S. Dweck

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Implicit Theories"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n28>

Print pages: 43-62

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n28>

[p. 43 ↓]

Chapter 28: Implicit Theories

Carol S.Dweck

Abstract

My enduring interest has been in the implicit theories, or basic beliefs, that people use to understand their world and to guide their behavior. In my research, I have found that one type of belief about human nature – the belief that fundamental human attributes are fixed traits or that they are malleable qualities that can be developed – has profound consequences for how people function, how they relate to others, and what they achieve. In this chapter, I trace the development of my interest in implicit theories from my beginnings in animal learning during the social awakening of the 1960s and the emerging cognitive revolution. Even then, I rejected the false distinction between basic and applied research, the false separation of affect, cognition, and motivation into different areas of study, and the false boundaries among fields of psychology (e.g., individual differences and social psychology), and I show how this rebellious stance informed and is embodied in my work. I end by showing how implicit theory research is making inroads into closing achievement gaps, promoting intergroup relations and conflict resolution, fostering cultures of productivity, and encouraging health behaviors.

Introduction

My abiding interest has been in the implicit theories, or basic beliefs, that people use to organize their world and to guide their behavior. I have been most fascinated by the fact that different people can form different basic beliefs. When one speaks of core knowledge about objects, space, time or number, psychologists assume that most people (unless or until they are trained in math or physics) achieve more or less the

same kind of understanding. However, when one considers basic beliefs about people and their attributes, different plausible positions are possible.

I have been particularly interested in beliefs with strong motivational properties. It might be interesting from an intellectual standpoint that people can come to different conclusions about the nature of themselves and others, but it becomes even more intriguing if the different conclusions make a difference for the goals people pursue and the outcomes they experience in their lives.

For some years I have studied the consequences of believing either that fundamental human attributes are fixed traits or that they are malleable qualities that can be developed (see Dweck, 1999, 2006). My collaborators and I have built and tested a model of the motivational, cognitive, affective, and behavioral consequences of the different implicit theories – and we have shown that these [p. 44 ↓] theories make a difference for people's achievement, relationships, careers, as well as their intergroup attitudes.

Personal Narrative and Intellectual History

I started my research career in a rat lab, studying animal learning, but the call of the cognitive revolution was too great. I had gone to graduate school at Yale to study animal learning, and the work was interesting, especially because I was in on the ground floor of the Wagner–Rescorla theory and because the work combined my interest in motivation and coping (e.g., Dweck and Wagner, 1970). The animal work on learned helplessness was also being conducted at that moment (Seligman, Maier, and Solomon., 1968). It had profound implications for how animals perceived reward contingencies, and how they used these perceptions to cope.

However, attribution theory was emerging and, for me, it held the promise of revealing how people interpreted the things that happened to them and how these interpretations guided the way in which they reacted. I could use my training in animal learning – training in parsimonious thinking and economical experimental design – and bring it to bear on the question of how people cope with the events that befall them.

Combining the seminal work on attribution theory (Weiner and Kukla, 1970) with the seminal work on learned helplessness (Seligman et al., 1968), I began to study how children coped with failure. My work revealed that children who attributed their failures to uncontrollable factors (e.g., their own lack of ability) showed a more helpless response to failure than those who attributed their failures to more controllable factors (e.g., their own effort) (Dweck and Reppucci, 1973). This helpless response to failure consisted of negative affect, falling expectancies, less effective strategies, and lower persistence, and did not in any way stem from lower ability.

I also provided evidence of a causal link between attributions and coping reactions through an intervention that changed children's attributions for failure and, in doing so, changed their helpless reactions to failure (Dweck, 1975). In my work, I have sought from the beginning to go back and forth between the lab and the field. The advantages of laboratory work are clear. You have a thrilling degree of control over what happens and how you measure its effects. Yet, you always need field work to tell you whether what you've elegantly controlled and measured bears any resemblance to what happens in the real world to people who are not under your experimental spell.

Graduate school was a fabulous experience. The faculty at Yale made us feel that we could and would change the world, and the cognitive revolution gave us the tools to attempt just that. Coming out of a period of behaviorism, in which the contents of the mind were forbidden territory, it was exhilarating to study beliefs, perceptions, construals, processing strategies, and the like in all their glorious manifestations. The late 1960s were a time of liberation. It was a time that was besotted with the idea of construction and it spawned a generation that rejected the oversimplified, deterministic, one-size-fits all behaviorist theory, as it rejected the one-size-fits all social constraints of the 1950s.

The 1960s and 1970s witnessed the emergence not only of cognitive psychology, but also social cognition in social psychology, cognitive therapy in clinical psychology, and social–cognitive approaches to personality. However, as with any revolution, some of the good things were thrown out with the bad. Cognitions are in the head and much of social psychology remained trapped in the head, giving short shrift to motivation, affect, behavior, and real life. Psychology became so cognitive that the august series *The Nebraska Symposium on Motivation* attempted to drop “motivation” from its title.

For me, however, the cognitive revolution meant that I could now address important [p. 45 ↓] outcomes with more tools at my disposal. I could now study how cognitive, motivational, and affective processes worked to produce behavior.

I also did not accept the idea that individual differences were not the domain of social psychology, or that social and personality psychology were essentially different fields. Virtually every individual difference I have ever measured, I have also induced experimentally. Both individual-difference measures and the experimental induction of beliefs are ways of understanding what makes people tick, of gaining leverage into the workings of the mind. Moreover, this combination of measurement and experimental induction captures the dynamic way in which people function. People may have strong and lasting beliefs, but they can also be swayed by a powerful situational cue or message.

In truth, I never accepted the idea of disciplinary boundaries within psychology at all. For convenience, psychologists have carved up the person into different parts – the cognitive part, the affective part, the social part, the developing part. This allows us to bring order to academic departments, journals, and organizations. But we should not be deceived into thinking that these boundaries are real. What we are seeking as researchers is an understanding of universal psychological processes and as we achieve this understanding we illuminate all areas of psychology. A commonly expressed fear is that neuroscience, as it burgeons, will reify these boundaries and make psychologists all the more parochial. My secret hope, however, is that neuroscience will do the opposite. The brain will not observe the boundaries psychologists have created, and will show instead how basic processes create commonality among disciplines.

Finally, I did not accept the idea, prevalent at the time, that in order to be scientific a psychological researcher had to avoid applied issues. The 1960s were, above all, a time when people cared about social issues and when unprecedented numbers of people became politically active. Ironically, as this was occurring, much of psychology was becoming increasingly abstract and “irrelevant.” Fortunately, Yale was one of the places where modern social psychology was born, as psychologists returned from World War II and tried to capture in their research phenomena like persuasion or obedience to

authority, phenomena that had played a role in the war. My mentors, fortunately, valued keeping a foot in the real world and making a difference.

My first job was at the University of Illinois, a wonderful department and, above all, an extraordinarily nurturing one. It is a place where people thrive. With my first graduate students, I took my work to the next level. We showed how the learned helplessness analysis (attribution processes) could shed light on gender differences in motivation and achievement. We demonstrated how girls, through their *better* treatment in grade school, could learn attributions for success and failure that would not serve them well later on when material became more difficult and success more uncertain (Dweck et al., 1978). We also modeled this process experimentally. Later, with Barbara Licht, we showed how girls' differing attributions for setbacks could illuminate their lower representation and achievement in math (Licht and Dweck, 1984). Here, we found the first evidence that the brightest girls might also be the most vulnerable. That is, we found a negative relation between IQ and girls' performance after a setback: the higher a girl's IQ, the less likely she was to master the material after a short period of confusion.

With Therese Goetz, we also showed that the helplessness model applied to social situations, and that we could predict via attributions who would show a helpless versus mastery-oriented response to social setbacks (Goetz and Dweck, 1980). Although many of these processes are easier to study in achievement/problem-solving situations, it has always been important for us to show that our model applies more broadly, and we have done so in every phase of the development of our implicit-theory model.

[p. 46 ↓]

With Carol Diener, we then fleshed out the helpless versus mastery-oriented responses to failure, monitoring the online moment-to-moment changes in cognition, affect, and behavior as children went from success to failure (Diener and Dweck, 1978). We learned many things from this work. What really hit home was the fact that different children were living in different psychological worlds. First, we saw how some children became excited and energized by difficulty. They were not simply "not helpless," but rather actively welcomed the challenge. Moreover, these were not students who had necessarily done better than others in the initial, success phase. What was also

interesting was that, unlike children who showed a helpless response, they did not seem to dwell on their difficulty, the reasons for it, and what it meant about them. In their talk-aloud narrative, they hardly ever even voiced attributions. Instead, they quickly became focused on mastering the new, more difficult problems. Finally, we monitored the exact problem-solving strategies students used and saw the helpless children as a group dissolved into ineffective strategies, while the mastery-oriented children remained highly strategic and taught themselves new ways to solve the problems. One child's self-negating failure was the other child's opportunity to learn. There seemed to be so much more than a simple difference in attribution involved. What else was going on?

An important piece of this puzzle emerged in my collaboration with Elaine Elliott and John Nicholls. In the course of intense discussions of achievement motivation over a period of time, we realized that achievement striving could be motivated by different goals: people could seek to demonstrate their ability (a *performance goal*) and/or to develop their ability (a *learning goal*). Elaine Elliott and I also realized that these different goals could be generating the starkly different helpless versus mastery-oriented responses we had observed in previous research. In research designed to test this hypothesis (Elliott and Dweck, 1988), our hunch was confirmed. When participants were led to hold strong performance goals and lost confidence in their abilities, we saw the whole helpless pattern of cognition, affect, and behavior emerge. Only when participants given performance goals were able to maintain high confidence in their abilities were they able to hold on and persist in the face of setbacks. In contrast, when participants were led to hold strong learning goals, they displayed a mastery-oriented response to setbacks – interestingly, even when they had low confidence in their ability. When the goal is to learn, one doesn't need to feel that one is already high in ability in order to remain engaged and persistent.

This achievement goal framework has generated a great deal of research, shedding new light on achievement processes and on academic outcomes in the real world. For example, in relatively recent work (Grant and Dweck, 2003), Heidi Grant examined students who were taking a highly challenging premed organic chemistry course, and showed that learning goals predicted the maintenance of intrinsic interest in the face of an initial poor grade, recovery from an initial poor grade, and higher final grades in the course. Performance goals (the desire to show high ability), on the other hand, predicted loss of intrinsic interest after an initial poor grade, a failure to recover from

an initial poor grade, and lower final grades in the course. Mediation analyses showed that learning goals predicted higher grades via deeper study strategies (see also Elliot, McGregor, and Gable, 1999) and via motivation-relevant self-regulation (e.g., keeping up interest in the subject matter).

The achievement goal analysis has also been successfully extended to organizational settings, to sports, and to issues in clinical psychology, for example, using chronic goal orientation as a predictor of depression (Dykman, 1998).

As generative as the achievement goal framework seemed to be, for me the picture was not complete. I still wondered why it would be the case that people with equal [p. 47 ↓] competence would chronically value and pursue different goals. Why were some people so concerned with proving over and over again how competent they were, whereas others were eagerly looking for challenges and opportunities to learn?

The next eureka moment came in a series of meetings with Mary Bandura. We suddenly realized that the idea of *ability* itself had a very different meaning when one thought about measuring and judging it through performance goals than when one thought about increasing it through learning goals. In the first case, *ability* connotes something deep-seated and permanent, whereas in the second case, *ability* implies something more dynamic and malleable. We then realized that these different conceptions of ability might lie behind differences in people's chronic goal choices. It was this hypothesis – that theories of intelligence would predict people's goal orientation – that was tested and supported in Mary Bandura's dissertation research.

Over the next few years, my students and I began to explore the ramifications of these implicit theories for motivation and behavior. Most memorably, Ellen Leggett and I spent day after day for several years developing the ideas into a broader motivational model, understanding and researching new aspects of the model, and developing implications of the model for personality as a whole (Dweck and Leggett, 1988). Little did I realize that many years later I would still be doing this!

This is not because I am so patient, focused, and systematic. It is because the model, for me, provides a microcosm of human functioning and thus operates on several levels. Aside from the immediate findings are the insights they can give into underlying

processes of cognition, affect, and behavior, and, at another level, the insights they can give into the nature of human personality, motivation, and dysfunction. I greatly admire psychologists who have used their more specific research to delve into basic human processes and to reflect on human nature (Mischel and Shoda, 1995; Bandura, 1986).

To best capture what the model has yielded to date, I will leave behind the chronology. Instead, I will describe the body of work (from my own and other laboratories) that has yielded a greater understanding of how implicit theories work, how they develop, how they affect important outcomes, and what role they serve in the larger scheme of human needs. At a more specific level, I will discuss the role they play in stereotyping, interpersonal interactions, group conflict resolution, and clinically relevant psychological processes. However, before doing so, I would like to underscore the importance of the exceptional colleagues I have been so fortunate to have in all the departments I have taught in. The atmosphere of passionate inquiry they fostered provided the perfect context for the development of ideas, and along with that, the development of enthusiastic and dedicated students. Those students, in turn, are the real stars of this research program.

Some Background Facts about Implicit Theories

What are the Entity and Incremental Theories?

The implicit theories are beliefs about the nature of human attributes. In the case of intelligence or of personality, for example, an *entity theorist* believes that the trait cannot be enhanced, whereas an *incremental theorist* believes that the trait can be developed. Those who hold an incremental theory do not necessarily believe that everyone starts out with the same talent or potential, or that anyone can be anything. They simply believe that everyone has the ability to grow with the proper motivation, opportunity, and instruction.

These are really beliefs about control, not stability. An entity theorist believes that people do not have control over their attributes or the power to change them. However, an entity theorist may believe that intelligence [p. 48 ↓] or personality can deteriorate with age. Moreover, an incremental theorist believes that people can change, but not necessarily that most people do change.

Throughout the chapter and throughout much of our research, we proceed as though people who endorse a given theory act consistently in terms of that theory, but the reality is bound to be more dynamic. That is, although the theories are found to be relatively stable across time (e.g., Robins and Pals, 2002), they can also be activated by strong cues or experiences in a situation (Good, Rattan, and Dweck, 2008; Murphy and Dweck, 2010).

Implicit theories are conceptually related to other variables, such as essentialist beliefs (e.g., Bastian and Haslam, 2006), beliefs about group ‘entitativity’ (e.g., Rydell et al., 2007), or beliefs about genetic determinism (e.g., Keller, 2005). All of these constructs capture the extent to which people or groups are seen as having deep-seated, somewhat immutable natures or structures, and the findings from these different lines of research are consistent with each other (Levy, Chiu, and Hong, 2006). The approach is also related to research on worldviews (e.g., Major et al., 2007; see Plaks, Grant, and Dweck, 2005), which seeks to capture the beliefs people use to organize and predict events in their lives.

Measures and Manipulations

We assess implicit theories by asking participants to agree or disagree with a series of statements, half of which present an entity theory and half of which present an incremental theory. In the domain of intelligence, for example, an entity theory item asserts, “You have a certain amount of intelligence, and you can't really do much to change it,” whereas an incremental item states, “No matter how much intelligence you have, you can always change it quite a bit.” In the domain of personality, an entity theory is tapped by items like, “The kind of person you are, is something very basic about you and it can't be changed very much,” whereas an incremental theory is reflected in items like, “You can change even your most basic qualities.” Using these

measures, on average, about 40 percent of people endorse an entity theory, 40 percent endorse an incremental theory, and 20 percent do not consistently endorse either theory.

Researchers have also developed domain-specific measures of implicit theories, for example, theories about particular abilities or domains, such as mathematics ability (Good et al., 2008), negotiation skills (Kray and Haselhuhn, 2007), managerial and decision-making skills (Tabernero and Wood, 1999), emotion regulation (Tamir, John, Srivastava, and Gross, 2007), or relationships (e.g., Knee, 1998). Researchers have also developed measures that apply to the self versus others (Dweck, 1999) or to group characteristics rather than individual characteristics (Halperin et al., 2009; Rydell et al., 2007; Tong and Chang, 2008). In each case, the measure asks whether the object in question can be changed/developed or not, and often the more specific and targeted measures have better predictive power (Rydell et al., 2007).

Many researchers have manipulated implicit theories. This has been done by giving instructions that portray the skill or domain in question as inherent and fixed or as learnable (Kray and Haselhuhn, 2007; Martocchio, 1994; Kasimatis, Miller, and Marcussen, 1996), by presenting participants with a “scientific” article to read that portrays the skill or domain as either fixed or malleable (Hong et al., 1999; Chiu et al., 1997b; Kray and Haselhuhn, 2007), or, in more long-term interventions, by presenting a workshop that teaches the incremental theory (and then comparing the results to control groups that learn potentially useful but theory-irrelevant lessons) (Aronson, Fried, and Good, 2002; Blackwell, Trzesniewski, and Dweck, 2007; Good, Aronson, and Inzlicht, 2003; see also Heslin and Vandewalle, 2008).

[p. 49 ↓]

Which Theory is True?

Both the entity and the incremental theory of intelligence have had their enthusiastic proponents. The entity theory was defended in *The Bell Curve* (Herrnstein and Murray, 1996), while the incremental theory was propounded forcefully by Alfred Binet (1909), the inventor of the IQ test, as well as the research sociologist Benjamin Bloom (1985),

paleontologist Steven Gould (1996), and creativity researcher John Hayes (1989). However, although both theories may have some truth, recent research by cognitive psychologists and neuroscientists is suggesting that fundamental aspects of executive function and intelligence can be taught not only in young children (Rueda et al., 2005), but also in college students (Jaeggi et al., 2008). In a study with college students (Jaeggi, et al., 2008), participants who were given training on a demanding working memory task, later scored significantly higher on an unrelated test of fluid intelligence. Fluid intelligence reflects the ability to reason and solve new problems. Moreover, the greater the training, the greater were the gains.

In the domain of personality too, researchers are reporting that even basic traits can show considerable change in adulthood (Roberts et al., 2006). In addition, as I have argued elsewhere, beliefs and belief systems themselves form a central part of personality that can be changed with targeted interventions, leading to widespread effects (Dweck, 2008).

When Do Implicit Theories Have the Strongest Effects?

In general, we find that implicit theories have the greatest effect when people are confronted with challenges or setbacks. For example, in a study of students making the difficult transition to seventh grade (Blackwell et al., 2007), entity and incremental theorists had shown no differences in prior math achievement in the more nurturing setting of elementary school; however, they showed a clear and continuing divergence in grades in their new, more challenging environment. In a related vein, we found that in a college calculus course or in a pre-med organic chemistry course, entity and incremental students showed diverging grades in these difficult courses as a function of their theories (Good et al., 2008) or their goals (Grant and Dweck, 2003).

Implicit theories also predict how people will judge other people, as I will describe below. People who believe in fixed traits engage in fundamentally different person judgment processes than do people who believe in malleable human qualities. However, experience has taught us that this only holds when people believe they are

forming and reporting their personal impressions of people, and not when they think they are performing a cognitive task with a right or wrong answer. When people are treating person information like variables in an equation that they are required to solve, their implicit theories play less of a role.

What Psychological Functions Do Implicit Theories Serve?

Implicit theories are beliefs about what people are made of and, by implication, how they work. As such, they should give people confidence that they can predict and control their social worlds. The work of Jason Plaks and his colleagues (Plaks, Grant, and Dweck, 2005; and Plaks and Stecher, 2007) provides evidence for this idea. They showed that when the predictions derived from people's implicit theories are violated, people experience anxiety and take steps to regain their sense of control. (Interestingly, this means that people will allow researchers to give them a new theory, as is done in experimental inductions or interventions, but they do not want to be left theory-less, that is, without a way of organizing and understanding how things work.)

[p. 50 ↓]

Meaning Systems: How Implicit Theories Work

Implicit theories create psychological worlds. They operate by recruiting allied goals and beliefs that work together as a “meaning system” (Molden and Dweck, 2006). These psychological worlds are portrayed below.

Goals

First, as I outlined earlier, the two implicit theories orient people toward different goals. Of course, everyone pursues all kinds of goals depending on the situation. Nonetheless, for people holding the entity theory, motivation tends to be organized more around validating their fixed traits via performance goals, whereas for people holding the incremental theory, motivation tends to be organized more around enhancing their malleable traits via learning goals (Beer, 2002; Dweck and Leggett, 1988; Kray and Haselhuhn, 2007; Mangels et al., 2006; Robins and Pals, 2002). Several studies have dramatically shown the lengths that people holding an entity theory of intelligence will go in order to look smart and not look dumb, often at the expense of important learning. For example, Hong et al. (1999) demonstrated that entity theorists express significantly less interest than incremental theorists in a remedial English course even when their English is poor and English proficiency is crucial to their academic success in college.

However, perhaps the most dramatic demonstration of the different goal orientations comes from an ERP (event-related potential) study, in which college students' brain waves were monitored for their patterns of attention as they took a very challenging general information test (Dweck et al., 2004; Mangels et al., 2006). Analysis of the brain-wave data showed that students who held an entity theory of their intelligence entered a strong state of attention to find out, after each question, whether they were right or wrong (satisfying a performance goal), but not to find out what the right answer really was, even when their answer had been wrong. In contrast, students who held an incremental theory of intelligence entered a strong state of attention both to find out whether their answer was correct (since that is also an important part of learning) and then again to find out what the correct answer really was. Indeed, when we later retested students on the questions they had missed (Mangels et al., 2006), these incremental students scored significantly higher on the retest than did those with the entity theory. Thus, different implicit theories appear to consistently engender different goals.

Effort Beliefs

According to attribution theory, effort is a controllable factor, and therefore the attribution of outcomes to effort should generate high motivation and resilience and, in general, this seems to be true. However, in the entity theory meaning system, there's a hitch: effort has negative implications for ability (Blackwell et al., 2007; Hong et al., 1999; Dweck and Leggett, 1988; Miele and Molden, 2009) – and ability is what entity theorists care about. In fact, working hard appears to quickly undermine entity theorists' confidence in their abilities. In a recent series of studies, Miele and Molden (2009) showed that any manipulation that gave participants a feeling of exerting effort (even something like small font size in a reading comprehension task) lowered entity theorists', but not incremental theorists', estimates of their ability/performance.

On the other hand, those with an incremental theory believe that high effort is good and necessary for the development of ability, and that even people who are geniuses have to work hard for their discoveries (Blackwell, et al., 2007; Dweck and Leggett, 1988). Their belief, by the way, is receiving increasing support, for example, in the work of Anders Ericsson (Ericsson et al., 1993), who finds that the most successful people in [p. 51 ↓] their fields are those who have engaged in the most deliberate practice and not necessarily those who seemed the most talented earlier on.

Incidentally, like the other variables in the meaning system, effort beliefs are not only correlated with implicit theories (e.g., implicit theories and effort beliefs showed a 0.54 correlation in a recent study of 373 adolescents; Blackwell et al., 2007) but also follow on the heel of an implicit theory induction (Hong et al., 1999)

Attributions

Implicit theories predict and generate different attributions for setbacks, with an entity theory orienting people more toward trait and ability attributions and an incremental theory orienting people more toward attributions that focus on effort or motivation. Our model does not argue that attributions are unimportant. Indeed several studies have shown them to be a key pathway from implicit theories to affective and behavioral

responses in challenging situations (Blackwell et al., 2007; Hong et al., 1999; Robins and Pals, 2002). However, attributions occur in the context of implicit theories and goals. For example, in a study that tracked students over their college years, Robins and Pals (2002) found that attributions were significantly predicted by implicit theories, both directly and indirectly through goals. In addition, when implicit theories are induced, the allied attributions tend to follow (e.g., Hong et al., 1999).

Helpless and Mastery-Oriented Strategies

The final link in the system, and the one that leads directly to important outcomes, consists of the different strategies that are fostered by the two implicit theories. Whereas the entity theory tends to lead to helpless or defensive strategies, the incremental theory fosters more persistent, strategic, mastery-oriented strategies. In experimental studies (e.g., Hong et al., 1999; Nussbaum and Dweck, 2008), those taught an entity theory of intelligence more often failed to confront their deficiencies and take steps to remedy them. Nussbaum and Dweck (2008) showed that after failure on a test, college students who were taught an entity theory of intelligence chose to repair their self-esteem, not through learning but through downward social comparison, that is, by looking at the tests of people who had done even worse. Those given an incremental theory overwhelmingly chose to learn by examining the tests of those who had done substantially better than they had. Nussbaum and Dweck also found that engineering students who had been given an entity theory did not choose to take a tutorial on the section of an engineering test on which they had done poorly, whereas those given an incremental theory overwhelmingly did so. In two longitudinal studies (Robins and Pals, 2002; Blackwell et al., 2007), students holding an entity theory were more likely than those holding an entity theory to report responding to academic difficulty with withdrawal of effort or cheating, and less likely to respond with new strategies or renewed effort.

Typically mediated by these strategy differences, the implicit theories have been shown to predict differences in key outcomes, such as grades (Blackwell et al., 2007), IQ test scores (Cury et al., 2008), changes in self-esteem over time (Robins and Pals, 2002), and negotiation success (Kray and Haselhun, 2007). In addition, as will be seen, interventions that teach an incremental theory yield improved outcomes in these and other areas.

Social Interactions and Social Relationships

Do implicit theories work similarly in other domains, such as interpersonal relationships? Indeed, implicit theories have been found to [p. 52 ↓] play a role in intimate relationships (Finkel et al., 2007; Kammrath and Dweck, 2006; Knee, 1998) and peer relationships in both children (Erdley et al., 1997) and adults (Beer, 2002).

Beer (2002), for example, showed that shy people who endorsed an incremental theory about their shyness ("I can change aspects of my shyness if I want to") elected to enter more challenging social situations, were more direct and active versus avoidant in their social interactions, and fared considerably better over the course of a new social interaction than did shy people who endorsed an entity theory about their shyness ("My shyness is something about me that I can't change very much").

In studies of intimate relationships, Ruvolo and Rotondo (1998) and Kammrath and Dweck (2006) measured participants' theories about the malleability of other people's personality ("The kind of person someone is, is something very basic about them and it can't be changed very much"), with the hypothesis that conflicts and setbacks would be more daunting when people believed their partners' flaws were permanent. And in fact, Ruvolo and Rotondo found that incremental theorists were better able to maintain relationship satisfaction even when they were faced with their partners' flaws or weaknesses. Further, Kammrath and Dweck found that following an important conflict, incremental theorists were more likely to give voice to their dissatisfaction in order to solve the problem. And, in several studies involving either romantic partners or peers, incremental theories were found to be more predictive of a tendency toward forgiveness as opposed to revenge (Finkel et al., 2007; Yeager et al., in press). Believing that others can change, it appears, allows people to take steps to influence them and work things out; believing that others cannot change leaves fewer good options: keep silent, leave, or seek payback.

Moreover implicit theories appear to operate in the social domain in similar ways to the intellectual-achievement domain. That is, people's self-theories are linked to their goals (Beer, 2002; Erdley et al., 1997; Knee, 1998), attributions (Erdley et al., 1997),

and mastery-oriented versus helpless responses to threat or setbacks (Beer, 2002; Kammrath and Dweck, 2006).

Person Perception, Social Judgment, and Stereotyping

It was not long before we began to ask whether implicit theories also affected how people perceived and judged others. If so, we might understand more about the basis of stereotyping and prejudice. This seemed especially interesting to us since it would mean that a belief that on the face of it had little or nothing to do with stereotyping could lay a foundation on which stereotypes thrived.

Ying-yi Hong, C.Y. Chiu, Cynthia Erdley, and I reasoned that the process of person judgment would be quite different for someone who believes that people are made up of fixed traits than for someone who believes that people are more dynamic and malleable. A belief in fixed traits should lead to a search for fixed traits, a relative neglect of the situation or the target's motivation, and more rigid judgments once they are rendered. This is exactly what we tested.

We found, first of all, that lay dispositionism and the fundamental attribution error were alive and well in entity theorists but were languishing in incremental theorists. For example, we found that entity theorists perceived almost any behavior as indicative of a person's underlying moral character (including such things as making one's bed in the morning) (Chiu et al., 1997b). Interestingly, they did not rate the behaviors themselves as better or worse than incremental theorists; they simply saw different implications for moral character. Entity theorists also more strongly believed that a person who was, say, more friendly or aggressive than another in one situation would also be [p. 53 ↓] more friendly or aggressive in a very different situation. Incremental theorists actually believed that the other guy would be the one to be more friendly or aggressive in the new and different situation – the opposite of the fundamental attribution error.

Next, we found that entity theorists were more likely to neglect salient information about the situation (Erdley and Dweck, 1993; Gervey et al., 1999; Molden et al., 2006;

Molden et al., 2006) or the target's motivation (Erdley & Dweck, 1993; Chiu, 1994) when making their judgments, but paid heightened attention to trait or trait-consistent information (Molden et al., 2006; Plaks et al., 2001). A similar bias was observed in their explanations for behavior: entity theorists were more likely to generate trait explanations for a target's behavior and less likely to think about psychological processes (e.g., motives, needs, construals) that could have caused the actions (Hong, 1994).

Moreover, even though entity theorists' trait inferences are drawn very rapidly, often from very preliminary information (Butler, 2000; Chiu et al., 1997b), they appear to have great confidence in them. They do not readily revise them in the face of counterinformation (Erdley and Dweck, 1993; Plaks et al., 2001) and they are willing to base decisions on them (Gervey et al., 1999). For example, Gervey et al. showed that entity theorists made strong inferences about moral character based on what the target, a defendant, was wearing on the day of the murder (a black leather jacket versus a business suit) and these judgments paralleled their guilty verdicts – to the point that potentially exonerating evidence had no impact on their decisions. And, believing they have judged a person as good or bad, entity theorists have a stronger tendency to endorse punishment as opposed to education for someone who has transgressed (Gervey et al., 1999; Erdley and Dweck, 1993; Chiu et al., 1997a).

Do these differences in person perception processes apply to the perception of groups and the formation of group stereotypes? Levy et al., (1998) and Levy and Dweck (1999) set out to explore this question by exposing participants to novel groups. Basically, people were given favorable or unfavorable information about some members of a group (or groups). We found that entity theorists formed stereotypes (global trait judgments of the groups) more readily, perceived greater homogeneity within groups and greater differences between groups, were more likely to generalize group traits to new members about whom they had no information, and had more extreme desire to interact or not interact with a group member based on the group information they had received.

We also found that entity theorists also had more stereotyped views of existing groups (Levy et al., 1998), and that they were more resistant to information that countered a stereotype (Plaks et al., 2001). In other words, as with the perception of individuals, those who held an incremental theory about human attributes made less extreme

and more provisional judgments that were open to revision. In fact, Plaks et al. found that incremental theorists were often more attentive to information that countered stereotypes than they were to information that confirmed them. Rydell et al. (2007) extended this work by examining people's theories about the fixed or malleable nature of groups (rather than individuals). Although they replicated past findings by showing that an entity theory about individuals predicted greater stereotyping, they also showed that an entity theory about groups – a more domain-specific measure – was an even better predictor of stereotyping.

For entity theorists there seems to be something "real" about belonging to a group, whether it is a social group, an occupational group, or a group based on race, ethnicity, or gender. For them, group members inevitably share traits. In a striking demonstration of this, Eberhardt et al., (2003) showed people pictures of biracial (morphed African-American and Caucasian) faces, telling them that a given face belonged either to an African-American or a Caucasian individual. [p. 54 ↓] When they were later asked to identify or draw the face, entity theorists chose/drew a face that accorded more with the label than did incremental theorists, who often chose/drew a face that moved farther away from the stereotype.

Yet, believing in fixed traits does not always predict greater stereotyping or prejudice. In very interesting work, Haslam and Levy (2006) showed that believing that gays' sexual preference was inborn and unchangeable predicted *less* prejudice. In this case, people apparently found it more acceptable to think of gays as having inborn tendencies than tendencies that were self-chosen and subject to personal change.

Person theories predict people's actual behavior toward groups as well. In studies of volunteering in the real world, Karafantis and Levy (2004) found that children's implicit person theories were related not only to their attitudes toward homeless children or poor children (e.g., their liking of them, desire to have contact with them, and their perceived similarity to them), but to their efforts on their behalf (volunteering, collecting money for UNICEF) and their enjoyment of those efforts.

Finally et al., (1998) showed that changing implicit theories changed people's tendency to form group stereotypes, along with their attitudes toward group members. Later, I describe a recent study that addressed the question of whether changing

implicit theories could change hardened intergroup attitudes and people's desire for reconciliation or compromise. We examined this in the context of the Arab-Israeli conflict, which brings us to our next topic.

Social Issues

Now that we have visited the two different worlds that implicit theories create, let us see whether this knowledge can illuminate social issues, such as longstanding group differences in achievement and intergroup relations. I will also ask whether implicit theories have a role to play in therapy and in issues of self-regulation and health.

Group Differences in Achievement

At the heart of American society is a desire for equal outcomes across groups. For this reason, differences between gender, racial, and ethnic groups in academic achievement are cause for great concern. Researchers therefore began to wonder whether implicit theories could shed light on processes that create these group differences and on interventions that can shrink these differences. When a negatively stereotyped person holds an entity theory (or believes that the people evaluating them do), one can see why they might be more vulnerable. In the face of difficulty, they may more readily think, "Maybe they're right. It's fixed and maybe I don't have it."

Thus experimental work has shown that when abilities are portrayed as fixed entities, stereotyped groups tend to show performance deficits on difficult tasks, but when the abilities are portrayed as experience-based or acquirable, those deficits are greatly attenuated or nonexistent. This has been shown for females and math (Dar-Nimrod and Heine, 2006), and for African-Americans in verbal areas (Aronson, 1998). In a similar vein, in a longitudinal study of college women in calculus, Good et al. (2008) found that when women perceived their math environment to portray math as a fixed ability, they were highly susceptible to stereotyping. In the face of stereotyping, they show a marked decrement in their sense that they belonged in math and, as they did, their desire to continue in math declined along with their course grades. However, when women perceived their math environment to be portraying math as an acquirable ability,

they were far less susceptible to stereotyping. Even when they reported high levels of stereotyping in their math environment, they were able to maintain a sense that they belonged in math, [p. 55 ↓] a desire to continue in math, and high grades in math.

Interestingly, when abilities are portrayed as fixed, the positively stereotyped group, such as men in math, can benefit (Mendoza-Denton et al., 2008). The idea that “it’s fixed and I have it” may indeed be motivating in the face of a difficult task. This fits well with findings by Reich and Arkin (2006), who showed that people are quite sensitive to the implicit theories that others hold about them. In this research, when participants were matched with evaluators who held an entity theory of their ability, they reported greater self-doubt when they expected to do poorly but less self-doubt when they expected to do well. Thus, an entity theory may increase the achievement gap both by depressing the confidence, motivation, and performance of the negatively stereotyped groups and by giving a boost to the positively stereotyped group.

Three implicit theory-based intervention studies have been conducted in academic settings, all showing an increase in motivation, grades, and/or achievement test scores for the experimental versus control groups (Aronson et al., 2002; Blackwell et al., 2007; Good et al., 2003). In these studies, middle school or college students in the experimental group learned an incremental theory of intelligence (that the brain forms new connections whenever they learn something new and that this learning makes them smarter over time) and how to apply this to their studies. Students in the control group learned other useful things, such as, in the Blackwell et al. study, a series of important study skills.

In the Aronson et al. (2002) study, African-American college students’ grades, enjoyment of academic work, and valuing of academic work increased significantly – even though their perceptions of negative stereotyping in their environment remained high. In the Good et al. (2003) study of adolescents, the gender difference in math performance, which was clear and significant in the control group, was greatly reduced and was not significant in the experimental groups. A similar pattern was found in a further analysis of the data from the Blackwell et al. (2007) study. Thus, the belief that abilities can be acquired, and messages to that effect from those in one’s learning environment, can help students fare better in challenging environments, and this appears to be especially so for targets of negative stereotypes.

Aside from direct interventions about the nature of ability, our research has shown that the type of praise students receive can have a striking effect on their implicit theories. This research was inspired by the self-esteem movement, with its gurus telling parents and educators to praise children's intelligence as lavishly and often as possible. Given our past findings, we thought this was bad advice. Sure enough, our studies (e.g., Mueller and Dweck, 1998) demonstrated that praise for intelligence (as opposed to praise for effort or strategy, i.e., *process praise*) encourages more of an entity theory and performance goals, and, in the face of difficulty, leads to greater decreases in motivation, confidence, and performance. Although this work has not directly addressed achievement gaps, it suggests that in trying to boost the confidence and achievement of underperforming groups, it would not be a good idea to praise their abilities. Rather, it suggests that focusing them on learning and on the processes that lead to success – effort, concentration, persistence, strategies – would be far preferable.

Intergroup Relations

Conflict Resolution

Because implicit theories appear to have far-reaching effects on attitudes toward other groups (Hong et al., 2004; Levy et al., 1998), they perhaps hold promise of reducing animosity and promoting accord between antagonistic groups. It may be an especially promising approach because changing implicit person theories does not involve directly trying to change people's attitude toward the "enemy," which would almost [p. 56 ↓] certainly meet with resistance. Rather it simply involves changing their ideas about people or groups in general. In new work, Halperin et al. (2009) show, first, that Israelis' attitudes towards peace with the Palestinians are predicted by their level of hatred for Palestinians; second, that implicit theories about groups predict Israelis' level of hatred toward Palestinians; and third, that fostering an incremental theory about groups in general both lowered Israelis' hatred of Palestinians and made the Israeli participants more favorable to a peace process. Inducing an incremental theory about groups was accomplished by means of an article that argued that groups do not have an inherent moral or immoral character but rather are incited to aggression by leaders and that

when the leaders change so may the group characteristics and behavior. No mention was made of Palestinians or their leaders.

Confronting Biased Behavior

Biased statements or actions present a good opportunity for educating outgroup members, particularly since such behavior is typically based on stereotypes or misinformation. However, confronting people and attempting to educate them presupposes that they can change. In new work, Rattan and Dweck (2010) show that when faced with biased remarks that included their group, people with incremental person theories were more likely to confront the speaker with the intent of educating him. Entity theorists, although they found the remark equally offensive, were not only less likely to confront the speaker, but also planned to avoid the speaker and people like him in the future. Additionally, we found that when people were led to hold an incremental person theory (by means of a scientific article that espoused and presented evidence for the theory), they were significantly more inclined to confront bias. Although not every situation permits the confronting of bias and although it is not incumbent upon negatively stereotyped individuals to confront bias whenever it arises, holding an incremental person theory may facilitate the process when it is appropriate or desirable.

Responses to Peer Bullying

Bullying and school violence have become a serious problems in schools around the world. I include this topic under intergroup relations because the victim of bullying is often a member of an outgroup, whether it is an ethnic or racial outgroup or a peer outgroup (e.g., computer nerds or kids who are physically different). Although the eradication of bullying is a top priority, it is also important to understand why some students respond to bullying with violent retaliation and others do not. Yeager et al. (in press), with sizable samples of high school students from the U.S. and Finland, either asked participants to recall a time when a peer had greatly upset them or gave them a vivid bullying scenario that was written as though it was happening to them. They were asked to choose the actions they would most feel like taking.

We found that implicit person theories predicted their preferred responses, with an entity theory consistently predicting festering resentment and the desire for violent, vengeful reactions (“hurting this person,” “imagining them getting hurt”). Moreover, an incremental theory intervention lessened students’ desire for violent revenge.

Management and Business

Now more than ever, business people must be responsive to the constant change that is taking place all around them, must be ready to correct the practices that are no longer working, and must be willing to try new approaches. To do otherwise is to risk stagnation or failure. Several lines of research have shown that implicit theories play a role in these processes. For example, Tabernero and Wood (1999) demonstrated the benefits of an incremental theory of management skills for the performance of individuals and work groups on challenging management tasks, in which new, corrective information was [p. 57 ↓] constantly being provided. Kray and Haselhuhn (2007) demonstrated that an implicit theory of negotiation ability predicted (and caused) superior negotiation outcomes, particularly on challenging tasks on which impasses were reached.

In an exciting program of research, Heslin and Vanderwalle (2008) showed that managers who held an entity person theory were less likely than their incremental counterparts (a) to be attuned to changes in employees’ performance after an initial good or poor performance was witnessed, remaining stuck in the initial impression; and (b) to mentor their employees, as reported by the employees themselves. Heslin and Vanderwalle then provided workshops that taught an incremental theory to a subset of the managers who had held entity theories. The managers who received this workshop, when tested six weeks later, displayed significantly more sensitivity to changes in an employee’s performance than did the managers in the control group, who had gone through a placebo workshop. In addition, they became more willing to provide mentorship and generated higher quality mentoring strategies. In summary, implicit theories have implications for learning, teaching, and productivity in a challenging, changing world.

Clinical Psychology, Psychotherapy, and Health

Because they affect self-regulation processes and interpersonal processes, implicit theories may well contribute to clinical psychology and to psychotherapy. First, research has shown that an entity theory and/or its allied goals (performance goals) play a role in depression (Dykman, 1998), in the loss of self-esteem following setbacks (Niiya et al., 2004), and in the negative impact of self-discrepancies (not matching one's ideal self) (Renaud and McConnell, 2007). There is also evidence that implicit theories about emotion regulation can play a role in emotional and social adjustment during the transition to college, with entity theorists experiencing waning social support and greater depression over time (Tamir et al., 2007).

Holding an entity theory of one's attributes increases defensiveness (Blackwell et al., 2007; Hong et al., 1999; Nussbaum and Dweck, 2008), which is a problem in itself, but can also greatly impede personal change, both within and outside of a therapeutic setting (see Dweck and Elliott-Moskwa, 2009, for a discussion of this). In addition, an incremental theory may predict better adherence to therapy, which inevitably is fraught with challenges and setbacks (see Dweck and Elliott-Moskwa, 2009, for a discussion of the potential role of implicit theories in cognitive behavior therapy). Preliminary evidence is also emerging to suggest that holding an incremental theory may predict better adherence to exercise and dieting in face of setbacks (Burnette, 2007; Kasimatis et al., 1996). This is an important area for future research and may well yield information about adherence to other health-maintaining or change-producing regimes.

Finally, therapists themselves may benefit from an incremental theory. Although most therapists, one hopes, hold the belief that people can change, they may approach very difficult patients (particularly ones who are threatening to their self-image as a competent therapist) with an entity theory. This may help protect the therapist from self-blame, but it may impede the therapeutic process if the therapist is less persistent in seeking strategies that can reach such clients (see Dweck and Elliott-Moskwa, 2009).

Conclusion

Research on implicit theories is giving us a portrait of people as dynamic creatures who are highly sensitive to cues in their environment and who are capable of change and growth. Moreover, the research is suggesting ways to promote that change and [p. 58 ↓] growth. As such, it is supporting a more incremental view of human abilities, human personality, and perhaps human nature.

When you begin a program of research, you have no idea where it will take you. I have stayed with this program of research because it continues to take me to new places. It remains challenging, it continues to yield provocative findings, and it has drawn me into the real world as people in the fields of education, business, sports, and health have sought to use our research to illuminate their practices. I cannot imagine a career more stimulating or more fulfilling – or one more conducive to personal growth.

References

- Aronson, J. (1998) The effects of conceiving ability as fixed or improvable on responses to stereotype threat . Unpublished manuscript, University of Texas.
- Aronson, J., Fried, C.B. and Good, C. Reducing the effects of stereotype threat on African American college students by shaping mindsets of intelligence . Journal of Experimental Social Psychology , 38, 113–125. (2002) <http://dx.doi.org/10.1006/jesp.2001.1491>
- Bandura, A. (1986) Social Foundations of Thought and Action: A Social Cognitive Theory . Englewood Cliffs, NJ: Prentice-Hall.
- Bastian, B. and Haslam, N. Psychological essentialism and stereotype endorsement . Journal of Experimental Social Psychology , 42, 228–235. (2006) <http://dx.doi.org/10.1016/j.jesp.2005.03.003>

Beer, J.S. Implicit self-theories of shyness . Journal of Personality and Social Psychology , 83, 1009–1024. (2002) <http://dx.doi.org/10.1037/0022-3514.83.4.1009>

Binet, A. (1909) Les idees modernes sur les enfants [Modern ideas on children] . Paris: Flammarion. (This edition 1973.)

Blackwell, L., Trzesniewski, K. and Dweck, C.S. Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention . Child Development , 78, 246–263. (2007) <http://dx.doi.org/10.1111/j.1467-8624.2007.00995.x>

Bloom, B.S. (1985) Developing Talent in Young People . New York: Ballantine.

Burnette, J.L. Implicit theories of weight management: A social cognitive approach to motivation . Dissertation Abstracts International , 67 (7-B) 4154. (2007)

Butler, R. Making judgments about ability: The role of implicit theories of ability . Journal of Personality and Social Psychology , 78, 965–978. (2000) <http://dx.doi.org/10.1037/0022-3514.78.5.965>

Chiu, C. (1994) Bases of categorization and person cognition . Unpublished PhD dissertation, Columbia University.

Chiu, C., Dweck, C.S., Tong, J.Y. and Fu, J.H. Implicit theories and conceptions of morality . Journal of Personality and Social Psychology , 73, 923–940. (1997a) <http://dx.doi.org/10.1037/0022-3514.73.5.923>

Chiu, C., Hong, Y. and Dweck, C.S. Lay dispositionism and implicit theories of personality . Journal of Personality and Social Psychology , 73, 19–30. (1997b) <http://dx.doi.org/10.1037/0022-3514.73.1.19>

Cury, F., Da Fonseca, D., Zahn, I. and Elliot, A. Implicit theories and IQ test performance: A sequential mediational analysis . Journal of Experimental Social Psychology , 44, 783–791. (2008) <http://dx.doi.org/10.1016/j.jesp.2007.07.003>

Dar-Nimrod, I. and Heine, S.J. Exposure to scientific theories affects women's math performance . *Science* , 314, 435. (2006) <http://dx.doi.org/10.1126/science.1131100>

Diener, C.I. and Dweck, C.S. An analysis of learned helplessness: Continuous changes in performance, strategy and achievement cognitions following failure . *Journal of Personality and Social Psychology* , 36, 451–462. (1978) <http://dx.doi.org/10.1037/0022-3514.36.5.451>

Dweck, C.S. The role of expectations and attributions in the alleviation of learned helplessness . *Journal of Personality and Social Psychology* , 31, 674–685. (1975) <http://dx.doi.org/10.1037/h0077149>

Dweck, C.S. (1999) Self-theories: Their Role in Motivation, Personality, and Development . Philadelphia: Psychology Press.

Dweck, C.S. (2006). Mindset . New York: Random House.

Dweck, C.S. Can personality be changed? The role of beliefs in personality and change . *Current Directions in Psychological Science* , 17, 391–394. (2008) <http://dx.doi.org/10.1111/j.1467-8721.2008.00612.x>

Dweck, C.S., Davidson, W., Nelson, S. and Enna, B. Sex differences in learned helplessness: (II) The contingencies of evaluative feedback in the classroom and (III) An experimental analysis . *Developmental Psychology* , 14, 268–276. (1978) <http://dx.doi.org/10.1037/0012-1649.14.3.268>

Dweck, C.S. and Elliott-Moskwa, E. (2009) Self-theories: The roots of defensiveness . In J.E. Maddux, ed. and J.P. Tangney (eds), *The Social Psychological Foundations of Clinical Psychology* . New York: Guilford Press.

Dweck, C.S. and Leggett, E.L. A social-cognitive approach to motivation and personality *Psychological Review* , 95, 256–273. (1988) <http://dx.doi.org/10.1037/0033-295X.95.2.256>

Dweck, C.S., Mangels, J. and Good, C. (2004) Motivational effects on attention, cognition, and performance . In D.Y. Dai, ed. and R.J. Sternberg (eds), *Motivation*,

Emotion, and Cognition: Integrated Perspectives on Intellectual Functioning . Mahwah, NJ: Erlbaum.

Dweck, C.S. and Reppucci, N.D. Learned helplessness and reinforcement responsibility in children . *Journal of Personality and Social Psychology* , 25, 109–116. (1973) <http://dx.doi.org/10.1037/h0034248>

Dweck, C.S. and Wagner, A.R. Situational cues and the correlation between CS and US as determinants of the conditioned emotional response . *Psychonomic Science* , 18, 145–147. (1970)

Dykman, B.M. Integrating cognitive and motivational factors in depression: Initial tests of a goal-oriented approach . *Journal of Personality and Social Psychology* , 74, 139–158. (1998) <http://dx.doi.org/10.1037/0022-3514.74.1.139>

Eberhardt, J.L., Dasgupta, N. and Banaszynski, T.L. Believing is seeing: The effects of racial labels and implicit beliefs on face perception . *Personality and Social Psychology Bulletin* , 29, 360–370. (2003) <http://dx.doi.org/10.1177/0146167202250215>

Elliot, A.J., McGregor, H.A. and Gable, S. Achievement goals, study strategies, and exam performance: A mediational analysis . *Journal of Educational Psychology* , 91, 549–563. (1999) <http://dx.doi.org/10.1037/0022-0663.91.3.549>

Elliott, E.S. and Dweck, C.S. Goals: An approach to motivation and achievement . *Journal of Personality and Social Psychology* , 54, 5–12. (1988) <http://dx.doi.org/10.1037/0022-3514.54.1.5>

Erdley, C., Cain, K., Loomis, C., Dumas-Hines, F. and Dweck, C.S. The relations among children's social goals, implicit personality theories and response to social failure . *Developmental Psychology* , 33, 263–272. (1997) <http://dx.doi.org/10.1037/0012-1649.33.2.263>

Erdley, C.S. and Dweck, C.S. Children's implicit theories as predictors of their social judgments . *Child Development* , 64, 863–878. (1993) <http://dx.doi.org/10.2307/1131223>

Ericsson, K.A., Krampe, R.T. and Tesch-Römer, C. The role of deliberate practice in the acquisition of expert performance . *Psychological Review* , 100, 363–406. (1993) <http://dx.doi.org/10.1037/0033-295X.100.3.363>

Finkel, E., BurnetteJ. and Scissors, L. Vengefully ever after: Destiny beliefs, state attachment anxiety, and forgiveness . *Journal of Personality and Social Psychology* , 92, 871–886. (2007) <http://dx.doi.org/10.1037/0022-3514.92.5.871>

Gervey, B., Chiu, C., Hong, Y. and Dweck, C.S. Differential use of person information in decision-making about guilt vs. innocence: The role of implicit theories . *Personality and Social Psychology Bulletin* , 25, 17–27. (1999) <http://dx.doi.org/10.1177/0146167299025001002>

Goetz, T.E. and Dweck, C.S. Learned helplessness in social situations . *Journal of Personality and Social Psychology* , 39, 246–255. (1980) <http://dx.doi.org/10.1037/0022-3514.39.2.246>

Good, C., Aronson, J. and Inzlicht, M. Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat . *Applied Developmental Psychology* , 24, 645–662. (2003) <http://dx.doi.org/10.1016/j.appdev.2003.09.002>

Good, C., Rattan, A. and Dweck, C.S. (2008) Development of the sense of belonging to math survey for adults: A longitudinal study of women in calculus . Unpublished manuscript.

Gould, S.J. (1996) *The Mismeasure of Man* . New York: Norton.

Grant, H. and Dweck, C.S. Clarifying achievement goals and their impact . *Journal of Personality and Social Psychology* , 85, 541–553. (2003) <http://dx.doi.org/10.1037/0022-3514.85.3.541>

Halperin, E., Russell, A., Dweck, C.S. and Gross, J. (2009) Emotion regulation in intergroup conflicts: Anger, hatred, implicit theories and prospects for peace . Unpublished manuscript.

Haslam, N. and Levy, S.R. Essentialist beliefs about homosexuality: Structure and implications for prejudice . *Personality and Social Psychology Bulletin* , 32, 471–485. (2006) <http://dx.doi.org/10.1177/0146167205276516>

Hayes, J.R. (1989) Cognitive processes in creativity . In J. Glover, ed. , R. Ronning, ed. , and C. Reynolds (eds), *Handbook of Creativity* . New York: Plenum.

Herrnstein, R. and Murray, C. (1996) *The Bell Curve* . New York: Simon & Schuster.

Heslin, P.A. and Vandewalle, D. Managers' implicit assumptions about personnel . *Current Directions in Psychological Science* , 17, 219–223. (2008) <http://dx.doi.org/10.1111/j.1467-8721.2008.00578.x>

Hong, Y. (1994) Predicting trait versus process inferences: The role of implicit theories . Unpublished PhD dissertation, Columbia University.

Hong, Y., Coleman, J., Chan, G., Wong, R.Y.M., Chiu, C., Hansen, I.G. Lee, S., Tong, Y. and Fu, H. Predicting intergroup bias: The interactive effects of implicit theory and social identity . *Personality and Social Psychology Bulletin* , 30, 1035–1047. (2004) <http://dx.doi.org/10.1177/0146167204264791>

Hong, Y.Y., Chiu, C., Dweck, C.S., Lin, D. and Wan, W. Implicit theories, attributions, and coping: A meaning system approach . *Journal of Personality and Social Psychology* , 77, 588–599. (1999) <http://dx.doi.org/10.1037/0022-3514.77.3.588>

Jaeggi, S.M., Buschkuhl, M., Jonides, J. and Perrig, W.J. Improving fluid intelligence with training on working memory . *Proceedings of the National Academy of Sciences* , 10, 14931–14936. (2008)

Kammrath, L. and Dweck, C.S. Voicing conflict: Preferred conflict strategies among incremental and entity theorists . *Personality and Social Psychology Bulletin* , 32, 1497–1508. (2006) <http://dx.doi.org/10.1177/0146167206291476>

Karafantis, D.M. and Levy, S.R. The role of children's lay theories about the malleability of human attributes in beliefs about and volunteering for disadvantaged groups . *Child Development* , 75, 236–250. (2004) <http://dx.doi.org/10.1111/j.1467-8624.2004.00666.x>

Kasimatis, M., Miller, M. and Marcussen, L. The effects of implicit theories on exercise motivation . *Journal of Research in Personality* , 30, 510–516. (1996) <http://dx.doi.org/10.1006/jrpe.1996.0037>

Keller, J. In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition . *Journal of Personality and Social Psychology* , 88, 686–702. (2005) <http://dx.doi.org/10.1037/0022-3514.88.4.686>

Knee, C.R. Implicit theories of relationships: Assessment and prediction of romantic relationship initiation, coping, and longevity . *Journal of Personality and Social Psychology* , 74, 360–370. (1998) <http://dx.doi.org/10.1037/0022-3514.74.2.360>

Kray, L.J. and Haselhuhn, M.P. Implicit negotiation beliefs and performance: Experimental and longitudinal evidence . *Journal of Personality and Social Psychology* , 93, 49–64. (2007) <http://dx.doi.org/10.1037/0022-3514.93.1.49>

Levy, S.R., Chiu, C. and Hong, Y. Lay theories and intergroup relations . *Group Processes and Intergroup Relations* , 9, 5–24. (2006) <http://dx.doi.org/10.1177/1368430206059855>

Levy, S.R. and Dweck, C.S. Children's static vs. dynamic person conceptions as predictors of their stereotype formation . *Child Development* , 70, 1163–1180. (1999) <http://dx.doi.org/10.1111/1467-8624.00085>

Levy, S., Stroessner, S. and Dweck, C.S. Stereotype formation and endorsement: The role of implicit theories . *Journal of Personality and Social Psychology* , 74, 1421–1436. (1998) <http://dx.doi.org/10.1037/0022-3514.74.6.1421>

Licht, B.G. and Dweck, C.S. Determinants of academic achievement: The interaction of children's achievement orientations with skill area . *Developmental Psychology* , 20, 628–636. (1984) <http://dx.doi.org/10.1037/0012-1649.20.4.628>

Major, B., Kaiser, C.R., O'Brien, L.T. and McCoy, S.K. Perceived discrimination as worldview threat or worldview confirmation: Implications for self-esteem . *Journal of Personality and Social Psychology* , 92, 1068–1086. (2007) <http://dx.doi.org/10.1037/0022-3514.92.6.1068>

Mangels, J.A., Butterfield, B., Lamb, J., Good, C.D. and Dweck, C.S. Why do beliefs about intelligence influence learning success? A social-cognitive-neuroscience model . *Social, Cognitive, and Affective Neuroscience* , 1, 75–86. (2006) <http://dx.doi.org/10.1093/scan/nsl013>

Martocchio, J.J. Effects of conceptions of ability on anxiety, self-efficacy, and learning in training . *Journal of Applied Psychology* , 79, 819–825. (1994) <http://dx.doi.org/10.1037/0021-9010.79.6.819>

Mendoza-Denton, R., Kahn, K. and Chan, W. Can fixed views of ability boost performance in the context of favorable stereotypes? *Journal of Experimental Social Psychology* , 44, 1187–1193. (2008) <http://dx.doi.org/10.1016/j.jesp.2008.03.005>

Miele, D.B. and Molden, D.C. (2009) Lay theories of intelligence and the role of processing fluency in perceived comprehension , Unpublished Manuscript, Northwestern University.

Mischel, W. and Shoda, Y. A cognitive-affective systems theory of personality: Reconceptualizing the invariances in personality and the role of situations . *Psychological Review* , 102, 246–268. (1995) <http://dx.doi.org/10.1037/0033-295X.102.2.246>

Molden, D.C. and Dweck, C.S. Finding ‘meaning’ in psychology: A lay theories approach to self-regulation, social perception, and social development . *American Psychologist* , 61, 192–203. (2006) <http://dx.doi.org/10.1037/0003-066X.61.3.192>

Molden, D.C., Plaks, J.E. and Dweck, C.S. ‘Meaningful’ social inferences: Effects of implicit theories on inferential processes . *Journal of Experimental Social Psychology* , 42, 738–752. (2006) <http://dx.doi.org/10.1016/j.jesp.2005.11.005>

Mueller, C.M. and Dweck, C.S. Intelligence praise can undermine motivation and performance . *Journal of Personality and Social Psychology* , 75, 33–52. (1998) <http://dx.doi.org/10.1037/0022-3514.75.1.33>

Murphy, M. and Dweck, C.S. A culture of genius: How an organization's lay theories shape people's cognition, affect, and behavior . *Personality and Social Psychology Bulletin* , 36, 283–96. (2010) <http://dx.doi.org/10.1177/0146167209347380>

Niiya, Y., Crocker, J. and Bartmess, E.N. From vulnerability to resilience: Learning orientations buffer contingent self-esteem from failure . *Psychological Science* , 15, 801–805. (2004) <http://dx.doi.org/10.1111/j.0956-7976.2004.00759.x>

Nussbaum, A.D. and Dweck, C.S. Defensiveness vs. remediation: Self-theories and modes of self-esteem maintenance . *Personality and Social Psychology Bulletin* , 34, 599–612. (2008) <http://dx.doi.org/10.1177/0146167207312960>

Plaks, J., Stroessner, S., Dweck, C.S. and Sherman, J. Person theories and attention allocation: Preference for stereotypic vs. counterstereotypic information . *Journal of Personality and Social Psychology* , 80, 876–893. (2001) <http://dx.doi.org/10.1037/0022-3514.80.6.876>

Plaks, J.E., Grant, H. and Dweck, C.S. Violations of implicit theories and the sense of prediction and control: Implications for motivated person perception . *Journal of Personality and Social Psychology* , 88, 245–262. (2005) <http://dx.doi.org/10.1037/0022-3514.88.2.245>

Plaks, J.E. and Stecher, K. Unexpected improvement, decline, and stasis: A prediction confidence perspective on achievement success and failure . *Journal of Personality and Social Psychology* , 93, 667–684. (2007) <http://dx.doi.org/10.1037/0022-3514.93.4.667>

Rattan, A. and Dweck, C.S. Who confronts prejudice? The role of implicit theories in the motivation to confront prejudice . *Psychological Science* , 21, 952–959. (2010) <http://dx.doi.org/10.1177/0956797610374740>

Reich, D.A. and Arkin, R.M. Self-doubt, attributions, and the perceived implicit theories of others . *Self and Identity* , 5, 89–109. (2006) <http://dx.doi.org/10.1080/15298860500441965>

Renaud, J.M. and McConnell, A.R. Wanting to be better but thinking you can't: Implicit theories of personality moderate the impact of self-discrepancies on self-esteem . *Self and Identity* , 6, 41–50. (2007) <http://dx.doi.org/10.1080/15298860600764597>

Roberts, B.W., Walton, K.E. and Viechtbauer, W. Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies . *Psychological Bulletin* , 132, 1–25. (2006) <http://dx.doi.org/10.1037/0033-2909.132.1.1>

Robins, R.W. and Pals, J.L. Implicit self-theories in the academic domain: Implications for goal orientation, attributions, affect, and self-esteem change . *Self and Identity* , 1, 313–336. (2002) <http://dx.doi.org/10.1080/15298860290106805>

Rueda, M.R., Rothbart, M.K., McCandliss, B.D., Saccomanno, L. and Posner, M.I. Training, maturation, and genetic influences on the development of executive attention . *Proceedings of the National Academy of Sciences* , 102, 14931–14936. (2005) <http://dx.doi.org/10.1073/pnas.0506897102>

Ruvolo, A.P. and Rotondo, J.L. Diamonds in the rough: Implicit personality theories and views of partner and self . *Personality and Social Psychology Bulletin* , 24, 750–758. (1998) <http://dx.doi.org/10.1177/0146167298247007>

Rydell, R.J., Hugenberg, K., Ray, D. and Mackie, D.M. Implicit theories about groups and stereotyping: The role of group entitativity . *Personality and Social Psychology Bulletin* , 33, 549–558. (2007) <http://dx.doi.org/10.1177/0146167206296956>

Seligman, M.E.P., Maier, S.F. and Geer, J.H. Alleviation of learned helplessness in the dog . *Journal of Abnormal Psychology* , 73, 256–272. (1968) <http://dx.doi.org/10.1037/h0025831>

Tabernero, C. and Wood, R.E. Implicit theories versus the social construal of ability in self-regulation and performance on a complex task . *Organizational Behavior and Human Decision Processes* , 78, 104–127. (1999) <http://dx.doi.org/10.1006/obhd.1999.2829>

Tamir, M., John, O.P., Srivastava, S. and Gross, J.J. Implicit theories of emotion: Affective and social outcomes across a major life transition . *Journal of Personality and Social Psychology* , 92, 731–744. (2007) <http://dx.doi.org/10.1037/0022-3514.92.4.731>

Tong, E. and Chang, W. Group entity belief: An individual difference construct based on implicit theories of social identities . *Journal of Personality* , 76, 707–732. (2008) <http://dx.doi.org/10.1111/j.1467-6494.2008.00501.x>

Weiner, B. and Kukla, A. An attributional analysis of achievement motivation . *Journal of Personality and Social Psychology* , 15, 1–20. (1970) <http://dx.doi.org/10.1037/h0029211>

Yeager, D.S., Trzesniewski, K.H., Tirri, K., Nokelainen, P. and Dweck, C.S. (in press) Adolescents' implicit theories predict desire for vengeance after peer conflicts: Correlational and experimental evidence . *Developmental Psychology* .

<http://dx.doi.org/10.4135/9781446249222.n28>

Handbook of Theories of Social Psychology

Uncertainty-Identity Theory

Contributors: Michael A. Hogg

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Uncertainty-Identity Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n29>

Print pages: 62-81

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n29>

[p. 62 ↓]

Chapter 29: Uncertainty-Identity Theory

Michael A.Hogg

Abstract

Uncertainty-identity theory is an account of how feelings of self-uncertainty motivate people to identify with groups in order to reduce their uncertainty. Self-uncertainty can very effectively be reduced by group identification because the process of self-categorization as a group member transforms self-conception so that it is governed by a group prototype that describes and prescribes what one should think, feel, and do, and how others will perceive and interact with you. It also provides consensual validation of self by fellow ingroup members. Group identification can make the world a more predictable place in which people know who they are, what to think, feel, and do, and how the course of interaction with others will play out. However, some groups, specifically high entitativity groups, have properties that make them better than others at reducing uncertainty. Uncertainty-identity theory has direct relevance to an explanation of a range of group phenomena, including social influence, norms, deviance, minority influence, schisms, leadership processes, and extremism and ideological orthodoxy.

Introduction

Social groups pervade the human experience. Language and symbolic communication, norms and culture, society and governance, commodities and the built environment – all are produced and configured by groups, and regulate and serve groups. Even our closest interpersonal and romantic relationships are shaped and contextualized by group membership. Ethnicity, religion, nation, organization, work team, sports club, family – these are all groups. They differ in size, distribution, distinctiveness, internal structure, longevity, purpose, “group-ness,” and of course what they do; but they all

share one fundamental characteristic – they provide their members with a sense of who they are, what they should think and do, and how others will perceive and treat them. Groups provide people with a social identity – a shared evaluation and definition of who one is and how one is located in the social world.

This perspective on groups as providers of social identity, and group behavior as a product of social identity processes, has been most fully explored over the past 40 years by social identity theory (Tajfel and Turner, 1979; Turner et al., 1987; also see Abrams and Hogg, 2010; Hogg, 2006; Hogg and Abrams, 1988). In the present chapter, I describe the development, concepts, and social relevance of *uncertainty-identity theory* (Hogg, 2007a), as an account of how [p. 63 ↓] social identity processes and phenomena are fundamentally motivated by people's need to reduce uncertainty about themselves. The theory has three basic premises: (1) people are motivated to reduce feelings of uncertainty about or related to themselves; (2) identifying with a group reduces self-uncertainty because the group's attributes are cognitively internalized as a *prototype* that describes and prescribes how one should behave and be treated by others, and one's prototype is consensually validated by fellow group members; and (3) highly entitative groups that are distinctive and clearly defined are most effective at reducing self-uncertainty.

Background and Development: A Personal Narrative

In the fall of 1978 I arrived at Bristol University to begin my PhD with John Turner. At that time Bristol was the uncontested center for social identity theory – Henri Tajfel was there, Rupert Brown had just left, and Marilyn Brewer was about to arrive for a sabbatical. The 1978 intake of PhD students also included Penny Oakes, Steven Reicher, and Margaret Wetherell. We were all working with John Turner to further develop the social identity concept with a more detailed focus on underlying cognitive mechanisms and on group processes in general. From this came the social identity theory of the group, self-categorization theory (Turner et al., 1987).

My own piece of the puzzle, for my dissertation, was group formation – what is the process through which a group forms or, more precisely, a person identifies with a group? At the time the prevalent social psychological model of group formation was group cohesiveness – interpersonal attraction is the social glue that holds groups together, and group formation is a matter, ultimately, of people developing bonds of attraction to one another. My doctoral research contested this model: distinguishing personal attraction, which has nothing to do with groups, from social attraction, which has everything to do with groups. Social attraction is the attraction based on categorization of oneself and others as members of the same group – group formation is a matter of self-categorization, not liking (Hogg, 1993).

Of course, the *process* of group formation as self-categorization (the “how”) invites the complementary question of what *motivates* people to identify with groups (the “why”). Although my dissertation focused on process, this motivational question interested me right from the start. At that time, the late 1970s, the main motivational dimension of social identity processes was positive intergroup distinctiveness and the pursuit of group-mediated self-esteem (e.g., Tajfel and Turner, 1979) – this aspect of social identity theory was critical as it described the mechanism for social change in intergroup relations. I quickly became involved with Turner in research on self-esteem and social identity.

In the fall of 1981, I became an assistant professor at Bristol and was joined in the fall of 1983 by Dominic Abrams. We had a productive (and very entertaining) 18 months together in Bristol before I headed to Sydney in early 1985 and then Melbourne in 1986, and Abrams headed to Dundee. In addition to planning and subsequently writing our social identity text, *Social Identifications* (Hogg and Abrams, 1988), we decided to clarify the motivational role of self-esteem in social identity processes. Our objective was to unpack and formalize what had already been written and systematize what the data actually showed. We coined the term *self-esteem hypothesis* and differentiated between two corollaries: low self-esteem motivates group identification, and group identification elevates self-esteem. We found that the evidence generally supported the latter but not the former corollary. Published in 1988 (Abrams and Hogg, 1988), this work was heavily cited, and its analysis and conclusions reiterated and supported in numerous subsequent papers by our colleagues.

Abrams and I wrote a number of follow-up chapters, and I can clearly recall sweltering [p. 64 ↓] in my study in Brisbane in the subtropical southern summer of 1991/1992, soon after having moved to the University of Queensland, finishing up yet another chapter on the self-esteem hypothesis. I was trying to add a novel twist and increasingly feeling that although positive distinctiveness and self-enhancement were important and might steer the course of group behavior there must be a more fundamental motivation for identifying with groups in the first place. It struck me that groups provide grounding for self-conception and social interaction – they reduce uncertainty about who one is – and that self-categorization is a cognitive process that is ideally suited to self-conceptual uncertainty-reduction. I rather sketchily added this idea to the end of the chapter and sent it to Abrams, telling him to delete it if he thought the idea made no sense. We went back and forth and decided this was actually quite a nice idea; so we retained it as a short three-page section at the end of the chapter and changed the title of the chapter accordingly – it appeared in 1993 (Hogg and Abrams, 1993).

The role of uncertainty as an epistemic motivation for social identity processes and group behaviors became my obsession and a main focus of my research. I gathered graduate students and postdoctoral fellows around me and we conducted empirical studies that all seemed to support the idea that people were more likely to identify with or would identify more strongly with a group if they were uncertain. I was now able to make a full conceptual statement, resting on data, of what at that time I still called the subjective uncertainty-reduction hypothesis – this was first published in preliminary form in 1999 (Hogg and Mullin, 1999) and then more completely in 2000 (Hogg, 2000).

It had struck me that some groups might be structured in ways that were better equipped to reduce uncertainty through group identification – entitativity seemed to fit the bill perfectly as a moderator of the uncertainty-identification relationship. Building on this idea I began to explore what would happen when uncertainty was subjectively more extreme or enduring – under these circumstances I argued that people would identify very strongly as zealots or true believers and that they would join groups that we would consider extremist, or make existing groups extremist. I had long been interested in how group membership can sometimes become all-enveloping – transforming people into zealous ideologues subscribing to inward-looking orthodoxies, treating dissenters and outsiders intolerantly as less than human, and endorsing and engaging in hurtful or brutal actions. These ideas were first published in 2004 (Hogg, 2004) and more fully

in 2005 (Hogg, 2005), and then with an application to autocratic corporate leadership structures in 2007 (Hogg, 2007b).

By now I had been using the term “uncertainty-identity theory” for a number of years and was ready to write a fully integrated and developed theoretical statement, which was published in 2007 (Hogg, 2007a). By mid 2006 when I moved to the Claremont Graduate University in Los Angeles, my interest in uncertainty-identity theory's potential as an explanation of group extremism had become a key focus of my own research and that of my students and collaborators. Building on my social identity theory of leadership (e.g., Hogg and van Knippenberg, 2003) a particular interest has been in the leadership dimension of uncertainty-sponsored extremism, and most recently in the possible role of active minorities and minority influence processes.

Basic Concepts, Processes, and Phenomena

The core tenet of uncertainty-identity theory is that feelings of uncertainty, particularly about or relating to who one is and how one should behave, motivate uncertainty-reduction, and that the process of self-categorization as a group member reduces self-conceptual [p. 65 ↓] uncertainty because it provides a consensually validated group prototype that describes and prescribes who one is and how one should behave.

Uncertainty

Feeling uncertain about one's perceptions, attitudes, feelings, and behaviors has a powerful motivational effect. We strive to reduce such uncertainties so that we feel less uncertain about the world we live in, and thus render it more predictable and our own behavior within it more efficacious. The American pragmatist philosopher John Dewey captures the motivational prominence of uncertainty-reduction rather nicely: “[I]n the absence of actual certainty in the midst of a precarious and hazardous world, men cultivate all sorts of things that would give them the *feeling of certainty*” (Dewey, 1929/2005: 33).

The experience of uncertainty can vary: it can be an exhilarating challenge to be confronted and resolved – it is exciting and makes us feel edgy and alive, and delivers us a sense of satisfaction and mastery when we resolve such uncertainties; or it can be anxiety provoking and stressful, making us feel impotent and unable to predict or control our world and what will happen to us in it. From the perspective of Blascovich's biopsychosocial model of challenge and threat (e.g., Blascovich and Tomaka, 1996) uncertainty can be thought of as a demand; if we believe our resources to deal with the demand are adequate we feel a sense of challenge that sponsors promotive or approach behaviors, if we believe our resources are inadequate we feel a sense of threat that sponsors more protective or avoidant behaviors. In this way people can reduce or regulate their uncertainty through quite different patterns of behavior that reflect a more promotive or more preventative approach (cf. Higgins's, 1998, regulatory focus theory).

The process of resolving uncertainty can be cognitively demanding. So in keeping with the cognitive miser or motivated tactician models of social cognition (e.g., Fiske and Taylor, 1991) we only expend cognitive energy resolving those uncertainties that are important or matter to us in a particular context. One of the key determinants of whether an uncertainty matters enough to warrant resolution is the extent to which self is involved. We are particularly motivated to reduce uncertainty if we feel uncertain about things that reflect on or are relevant to self, or if we are uncertain about self per se; about our identity, who we are, how we relate to others, and how we are socially located. Ultimately, people need to know who they are, how to behave, and what to think; and who others are, how they might behave, and what they might think.

In talking about uncertainty, it is more appropriate to talk about reducing uncertainty than achieving certainty – people cannot feel completely certain but only less uncertain (Pollock, 2003). Absolute certitude is generally viewed as a dangerous delusion that is the province of narcissists, zealots, ideologues, and true believers. People typically work to reduce uncertainty until they are “sufficiently” certain about something to desist from dedicating further cognitive effort to uncertainty reduction – this provides closure, in the gestalt sense (Koffka, 1935), and allows one to move on to dedicate cognitive effort to other things. Hence uncertainty-identity theory is about reducing uncertainty rather than achieving certainty.

However, there is a caveat. The pursuit of uncertainty reduction does not rule out the possibility that individuals or groups sometimes embark on courses of action that may in the short term increase uncertainty. Typically this might happen when an existing state of affairs is marked by glaring contradictions that engender uncertainty. This idea has parallels with the way that formal science progresses – periods of “normal science” where uncertainty is low and small contradictions accumulate but are concealed, punctuated by “scientific revolutions” where contradictions and uncertainties burst to the [p. 66 ↓] fore to sponsor a “paradigm shift” and subsequent reduction of uncertainty (Kuhn, 1962). Another example is when a current state of affairs in one's life or the society in which one lives is unbearable and a measured risk must be taken to improve things – change is risky and uncertain and therefore not undertaken lightly (e.g., Jost and Hunyady, 2002).

The idea that uncertainty plays a significant role in motivating human behavior is not new (Fromm, 1947), and there are many social psychological analyses of the causes and consequences of uncertainty (e.g., Kahneman et al., 1982). Because uncertainty-identity theory links self-uncertainty to identity and group processes, the treatments of uncertainty that are most relevant here are those that focus on social comparison processes (e.g., Festinger, 1954; Suls and Wheeler, 2000), uncertainty about or related to self (e.g., Arkin et al., 2010), individual and cultural differences in peoples tolerance and reaction to uncertainty (e.g., Hofstede, 1980; Kruglanski and Webster, 1996; Schwartz, 1992; Sorrentino and Roney, 1999), communication to reduce uncertainty (e.g., Berger, 1987), uncertainty and organizational socialization (e.g., Saks and Ashforth, 1997), and uncertainty-sponsored zealotry and defense of one's world-view (e.g., Kruglanski et al., 2006; McGregor and Marigold, 2003; McGregor et al., 2001; Van den Bos, 2009; Van den Bos et al., 2005). How these ideas relate to uncertainty-identity theory is discussed in detail elsewhere (Hogg, 2007a, 2010a), but the key differences are that (a) social identity and collective self are center-stage, (b) uncertainty is context-not personality-dependent, (c) a social cognitive process transforms uncertainty into group behavior, and (d) it is a motivational account of group phenomena in general not just extremism.

Uncertainty is related to meaning, and some argue that the primary human motive is a search for meaning (e.g., Bartlett, 1932; Maslow, 1987) – people strive to construct a coherent and meaningful worldview. This idea has been explored recently by Heine

et al., (2006) in their meaning maintenance model (cf. Swann's self-verification theory [Swann et al., 2003]): people are meaning-makers driven to establish associative frameworks that (a) tie together elements of the world, (b) tie together elements of themselves, and (c) most importantly bind self to the world. This view of meaning focuses less on meaning than on associative links that one has confidence in and one is certain about, suggesting that reduced uncertainty is critical and may motivationally underpin meaning. From an uncertainty-identity perspective uncertainty reduction is closely linked to meaning-making, however it is the aversive feeling of uncertainty that actually motivates. For example, baseball makes no sense to most Italians and yet they have little motivation to make it meaningful – they are uncertain but do not care as they do not feel it is important to who they are that the game should make sense to them. For Italians, soccer is an entirely different motivational matter.

Finally, uncertainty takes many forms and has many foci. It can be wide ranging and diffuse; for example, feeling uncertain about one's future, or very specific and focused, for example feeling uncertain about what to wear to a party. Feelings of uncertainty can vary in strength and be transitory or enduring – however, from an uncertainty-identity theory, perspective enduring uncertainty is not primarily a matter of personality but rather a reflection of an enduring context that creates uncertainty. As mentioned earlier, uncertainty also varies in the degree to which it reflects on or relates to self-conception in a particular context. Uncertainty about or related to self is likely to have the greatest motivational force because the self is the critical organizing principle, referent point, or integrative framework for perceptions, feelings, and behaviors. It is this self-uncertainty that is most directly implicated in social identity processes.

[p. 67 ↓]

Self-Categorization and Group Identification

Feelings of uncertainty have different causes and different foci. Uncertainty-identity theory focuses on context-induced feelings of uncertainty that are about self or about things that relate to, reflect on, or matter to self. If a particular context that induces uncertainty endures, for example a long-lasting economic crisis, uncertainty and attempts to reduce or fend off uncertainty may also endure. There may be individual

differences in how much uncertainty people feel in a given context and in how people respond to uncertainty; however, this is treated, to use a statistical metaphor, as error variance – it is not the focus of uncertainty-identity theory. This orientation toward personality and individual differences is consistent with the group-focused metatheory that informs social identity theory and uncertainty-identity theory (e.g., Abrams and Hogg, 2004; Turner, 1999; also see Hogg, 2008).

Feelings of uncertainty about or reflecting on self can be resolved in many different ways – for example, one can introspect. However, the crux of uncertainty-identity theory is that group identification, via the process of self-categorization (e.g., Turner et al., 1987), is one of the most powerful and effective ways to reduce self-uncertainty. Human groups are social categories that we cognitively represent as *prototypes* – fuzzy sets of attributes (e.g., perceptions, beliefs, attitudes, values, feelings, behaviors) that define the category and distinguish it from other categories in a specific context. Prototypes describe behaviors; but the prototype of a group we belong to also has prescriptive properties that dictate how we *ought* to behave as a group member.

Prototypes obey the metacontrast principle – they embody attributes that maximize the ratio of intergroup differences to intragroup differences, and thus accentuate perceived differences between and similarities within groups (cf. Tajfel, 1959). This principle ensures that the prototype of a specific group is influenced by what group it is being compared to and for what purpose, and thus that group prototypes are not simply average group attributes but are often ideal group attributes.

When we categorize someone as a member of a specific group we assign the group's prototypical attributes to that person. We view them through the lens of the prototype of that group; seeing them not as unique individuals but as more or less prototypical group members – a process called *depersonalization*. When we categorize others, ingroup or outgroup members, we stereotype them and have expectations of what they think and feel and how they will behave. When we categorize ourselves, self-categorization, exactly the same process occurs – we assign prescriptive ingroup attributes to ourselves, we auto-stereotype, conform to group norms, and transform our self conception.

In this way group identification is very effective at reducing self-related uncertainty. It provides us with a sense of who we are that prescribes what we should think, feel, and do. Because self-categorization is inextricably linked to categorization of others, it also reduces uncertainty about how others, both ingroup and outgroup members, will behave and what course social interaction will take. It also provides consensual validation of our worldview and sense of self, which further reduces uncertainty. Because people in a group tend to share the same prototype of “us” and share the same prototype of “them,” our own expectations about the prototype-based behavior of others are usually confirmed, and our fellow group members agree with our perceptions, beliefs, attitudes, and values and approve of how we behave. The discovery that fellow ingroup members do not see the world as we do can be a source of profound uncertainty about what the group stands for and thus about self-conception. I discuss this in the following text.

[p. 68 ↓]

Clearly, identification can effectively reduce uncertainty and protect one from uncertainty. The implication is that uncertainty reduction motivates group identification – we identify with groups in order to reduce, or protect ourselves from, uncertainty. When people feel uncertain about themselves or things reflecting on self they “join” new groups (e.g., sign up as a member of a community action group), identify with or identify more strongly with existing self-inclusive categories (e.g., one's nation), or identify with or identify more strongly with groups that they already “belong” to (e.g., one's work team).

Uncertainty reduction guides the process responsible for making a social categorization psychologically salient as the basis for self-categorization, group identification, and group behavior. People draw on accessible social categorizations – ones they value, find important, and frequently use to define themselves and perceive others (they are *chronically accessible* in memory), and/or ones that are prominent and self-evident in the immediate situation (they are *situationally accessible*) (e.g., Oakes, 1987; Turner et al., 1994) – gender is often chronically and situationally accessible. People then investigate how well the categorization accounts for similarities and differences among people (*structural/comparative fit* – do the men and women behave differently?) and how well the stereotypical properties of the categorization account for why people

behave as they do (*normative fit* – do the men and women behave in accordance with one's gender expectations/stereotypes?).

If the fit of a particular categorization is poor, people cycle through other accessible categorizations until an optimal fit is obtained. This process is primarily fast and automatic because people strive to reduce feelings of uncertainty about self-conception, social interaction, and people's behavior. The very notion that an accessible categorization needs to *fit* implies that it reduces feelings of uncertainty about the social context and our place within it. The categorization with optimal fit becomes psychologically salient as the basis of self-categorization, group identification, and prototype-based depersonalization. It triggers social identity-related perceptions, cognitions, affect, and behavior.

The uncertainty-identity theory conception of the relation between uncertainty and group identification represents a relatively hydraulic model of group motivation. Uncertainty, however induced, mobilizes one to psychologically identify and is reduced by identification. However, feelings of uncertainty are multiply determined and can be addressed in many different ways. Identification is only one way to address uncertainty, but one that is particularly effective in the case of self-related uncertainties. Feelings of uncertainty can also be fleeting. As soon as one uncertainty is reduced one's mind can be assailed by new uncertainties or we can seek out new ones to resolve.

The most basic prediction that can be made from uncertainty-identity theory is that the more uncertain people are the more likely they are to identify, and to identify more strongly, with a self-inclusive social category. This prediction has been confirmed across a number of minimal group studies in which people identified with relatively minimal groups and engaged in ingroup favoritism and intergroup discrimination only when they were categorized under uncertainty (e.g., Grieve and Hogg, 1999; Mullin and Hogg, 1998; for an overview of this research see Hogg, 2000, 2007a). In these studies uncertainty was manipulated in a variety of ways. For example, participants described what they thought was happening in ambiguous or unambiguous pictures, or they estimated the number of objects displayed in pictures in which there were very few objects or so many objects that they could only make a wild guess.

Other studies showed that uncertainty was a stronger motivation for identification if participants were uncertain about something they felt was important and self-relevant, and if the prototypical properties of the available [p. 69 ↓] social category were relevant to the focus of uncertainty. There are also studies showing that uncertainty significantly strengthens identification even when depressed self-esteem as a possible mediator is controlled for (Hogg and Svensson, 2010), and even when people are actually placed in a relatively low status group (Reid and Hogg, 2005).

Entitativity

Are there generic properties of groups that might make some types of groups and associated identities and prototypes better equipped to reduce uncertainty through identification? I have touched on a few of these potential moderators of the uncertainty-identity relationship above, but the most significant is entitativity (Hogg, 2004). Entitativity is that property of a group, resting on clear boundaries, internal homogeneity, social interaction, clear internal structure, common goals, and common fate, which makes a group “groupy” (Campbell, 1958; Hamilton and Sherman, 1996). Groups can vary widely in entitativity from a loose aggregate to a highly distinctive and cohesive unit (e.g., Lickel et al., 2000).

Group identification reduces uncertainty because it provides a clear sense of self that prescribes behavior and renders social interaction predictable. An unclearly structured low entitativity group that has indistinct boundaries, ambiguous membership criteria, limited shared goals, and little agreement on group attributes will do a poor job of reducing or fending off self-related uncertainty. In contrast a clearly structured high entitativity group with sharp boundaries, unambiguous membership criteria, tightly shared goals, and consensus on group attributes will do an excellent job. Identification via self-categorization reduces uncertainty because self is governed by a prototype that prescribes cognition, affect, and behavior. Prototypes that are simple, clear, unambiguous, prescriptive, focused, and consensual are more effective than those that are vague, ambiguous, unfocused, and dissensual. Clear prototypes, such as the former, are more likely to be grounded in high than low entitativity groups. In addition, people are more likely to anchor the attributes of high entitativity groups in invariant underlying qualities or essences (e.g., Haslam et al., 1998) that provide further

interpretative predictability and stability and make the group even better at reducing and fending off uncertainty.

From uncertainty-identity theory the prediction is that although under uncertainty, especially self-uncertainty, people will identify with groups, they will show a strong preference for high entitativity groups. People will seek out highly entitative groups with which to identify or they will work to elevate, subjectively or actually, the entitativity of groups to which they already belong. This idea has support from a number of indirect investigations of uncertainty, entitativity, and group identification (e.g., Castano et al., 2003; Jetten et al., 2000; Pickett and Brewer, 2001; Pickett et al., 2002; Yzerbyt et al., 2000).

Direct tests have provided better and more robust support (Hogg et al., 2007; Sherman et al., 2009). Hogg et al. (2007) conducted two studies in which uncertainty, explicitly about self, was experimentally primed to be high or low, and the perceived entitativity of participants' political party was measured (Study 1) or the perceived entitativity of participants' ad hoc lab group was manipulated (Study 2). Group identification was measured by a multi-item scale. In both cases, participants identified significantly more strongly when they were uncertain and their group was highly entitative. Sherman et al. (2009) conducted a pair of field studies of political party supporters and workers on strike, to provide support for the related idea that self-uncertainty can lead people to perceptually polarize groups in order to accentuate the perceived entitativity of a group they are already members of.

[p. 70 ↓]

Implications, Extensions, and Applications

Group identification-based reduction of self-uncertainty has a large array of implications and applications. These implications extend the conceptual reach of the theory to embrace social influence; leadership processes and trust; dissent, deviance, and minority influence; and extremism and ideological orthodoxy.

Self-Uncertainty and Group Influence

As described earlier, one of the ways that social identity reduces self-uncertainty is that it anchors self-conception in a consensual world view – it surrounds one with fellow ingroupers who see the world largely in the same way as you do and who thus provide consensual validation of your perceptions, attitudes, behaviors, and ultimately self-concept. Not surprisingly, discovering that fellow ingroupers do not see the world as you do can be a source of profound self-uncertainty. Indeed, disagreement with fellow ingroups (i.e., normative disagreement) often initiates a process of social influence, *referent informational influence*, through which people urgently seek information to confirm the group's norms and identity in order to know what they are identifying with and therefore what their own identity is (Turner, 1991; Turner et al. 1987; also see Hogg and Smith, 2007; McGarty et al., 1993).

Given how important it is for self-uncertainty management to know what the ingroup prototype/norm is, people can spend substantial time engaged in ‘norm talk’ – communicating mainly with fellow ingroup members in order to be sure of the groups defining and prescriptive attributes (Hogg and Reid, 2006). Through this normative communication process people tend to look to prototypical ingroup members to provide the most reliable information about ingroup norms. However, outgroups can also be a useful source of normative information. Discovering that you are in disagreement with an outgroup does not produce uncertainty; it is expected and serves to confirm your ingroup identity – they are what you are not. Agreement with an outgroup would, however, be problematic. It would produce self-related uncertainty, causing you to question what your group stands for and whether you really fit in as a member.

Leadership and Trust

Within groups, prototypical members are more influential than nonprototypical members and people pay closer attention to the former than the latter as a reliable source of information about group norms and social identity (Hogg, 2010 b). Furthermore, according to the social identity theory of leadership (Hogg, 2001; Hogg and van

Knippenberg, 2003) prototypical members tend to occupy leadership positions, and leaders are more effective if they are perceived to be prototypical. The clear implication is that self-uncertainty will cause people to pay even closer attention to their leaders, be more likely to empower them and follow them, and thirst for recognition and validation by them. Under these circumstances people will need to feel they are valued by and can trust their leaders (Lind and Tyler, 1988; Tyler, 1997; Tyler and Lind, 1992), even if such feelings are actually an illusion.

Trust plays an important role here – it is closely associated with predictability and uncertainty. The more we are able to trust someone the more predictable they are and the less uncertain we feel. Trust plays a central role in group, specifically ingroup, life. Ingroups are “bounded communities of mutual trust and obligation” (Brewer, 1999: 433) in which members expect to be able to trust fellow members to do them no harm and to be acting in the best interest of the group. Ingroup members who betray our group-based trust, by leaving the group to pursue their personal interest or by acting in ways [p. 71 ↓] that only benefit themselves and are to the detriment of the group as a whole, reduce trust and raise uncertainty and thus invite harsh reactions. This dynamic is particularly pronounced for central members who are prototypical of the group or act as group leaders. Disloyalty and violation of trust on the part of prototypical members is highly disruptive of group equilibrium (e.g., van Vugt and Hart, 2004), and is a particularly potent source of uncertainty about what the group stands for, about one's membership in the group, and ultimately about self.

Because uncertainty can empower leaders it would not be surprising to discover that smart leaders use uncertainty strategically as a resource. Marris's (1996) analysis of the politics of uncertainty supports this idea – certainty is power and the powerful can create uncertainty for the powerless in order to control them. Thus leaders, particularly those who feel the group perceives them to be prototypical and trusts them, can reinstate, maintain or strengthen their authority by engaging in a rhetoric of uncertainty coupled with reassurances that they are able to resolve the uncertainty. This would have the effect of strengthening followers' identification with the group and their support for the leader – a study by Hohman et al. (2010) provides some support for this idea. Furthermore, the media often has examples of national, religious, and corporate leaders doing precisely this – and it is a strategy that can, as I discuss below, produce autocratic or extremist leadership (Hogg, 2007b).

Outsiders: Marginal Members, Deviants, and Minorities

Leaders, as we have seen, play a significant role in resolving self-uncertainty because members look to them to define who “we” are – to resolve normative disagreement or ambiguity. However, normative disagreement and ambiguity within a group can produce an array of responses. It undermines the group's entitativity and thus one response is to disidentify or identify less strongly with the group and identify with other more entitative groups. The tendency for uncertainty to weaken identification with low entitativity groups has been empirically confirmed in a pair of studies of moderate and radical groups by Hogg, Farquharson et al. (2010).

This notion of normative ambiguity, disagreement, or conflict can be taken further by focusing on the source of disagreement. Where disagreement seems to be with the leader's normative example and the leader is trusted and prototypical, then, as we saw above, members' realign themselves with the leader and identify with the group. However, where the leader is less prototypical and/or less trusted, members may initiate or engineer leadership change, or they may simply feel they themselves no longer fit the group and therefore that the group does not anchor their identity, so they weaken their ties, disidentify, and leave the group. Where normative disagreement is with a less prototypical, marginal member who does not occupy a leadership role, little normative uncertainty is evoked, much like encountering disagreement with an outgroup member. On the contrary, normative divergence on the part of nonprototypical members motivates the group to pressure the deviate to conform, but also may invite derogation, marginalization, persecution, and ejection from the group, often orchestrated by the group's leadership (e.g., Hogg et al., 2005; Marques et al., 2001; Marques and Paez, 1994).

Marginal group members who express normative divergence can sometimes avoid derogation and instead play an active role in normative clarification and uncertainty reduction. Specifically this may happen when they do not simply diverge, but rather act as constructive ingroup critics oriented toward clarification and improvement of ingroup normative practices (e.g., Hornsey, 2005). The other way in which normative deviance

can impact the group is when those who diverge from normative practices and thus fracture entitativity and potentially raise [p. 72 ↓] uncertainty about the group and its identity are themselves a collective or group. This is very common, and more often than not creates a schism around alternative views of what the group stands for (e.g., Sani and Reicher, 1998). This state of affairs creates great uncertainty that usually requires members to identify strongly with one or other faction in order to reduce uncertainty, and of course the original group is usually changed forever.

Another way to view normatively divergent subgroups is as active minorities engaging in minority influence. Indeed, research on minority influence argues that active minorities who maintain their novel alternative position in a consistent manner ultimately change majority views quite radically precisely because they make the majority uncertain about the validity of the majority's position (e.g., Martin and Hewstone, 2008; Moscovici, 1980; Mugny and Pérez, 1991). The uncertainty is resolved by reconfiguring the majority position and identifying strongly with it, a conversion effect associated with social change. Much like it is easier to brush off disagreement with and criticism from outgroup members than fellow ingroupers, outgroup minorities may produce less uncertainty and subsequent normative change than ingroup minorities. This is consistent with the self-categorization theory analysis of minority influence (e.g., David and Turner, 2001) and with Crano's view that ingroup minorities can be more effective because a *leniency contract* is struck in which the majority agrees to be lenient toward the minority and pay attention to its views as long as the minority desists from being "too extreme" (e.g., Crano and Seyranian, 2009).

The discussion in this section is predicated on the idea that group identification effectively reduces self-uncertainty because the group prototype unambiguously defines and prescribes self, and because there is consensual validation of one's social identity. Intragroup disagreement and the existence of normative deviants potentially undermine entitativity and weaken the uncertainty reduction capacity of the group. We explored how different types of normative divergence may sponsor different responses.

A final aspect of this analysis concerns perceived self-prototypicality. What is important here for identity validation and thus uncertainty reduction is feeling that other members of the group believe you are prototypical, and accept and include you as a bone fide member who "fits in." Clearly, if you yourself feel you do not really fit in then the group,

however highly entitative it may be, will not very effectively resolve self-uncertainty – you will feel like an imposter, a square peg in a round hole. Indeed poor fit may be a more serious identification and uncertainty management issue in highly entitative than less entitative groups. Even where you yourself feel you fit or that you will work hard to fit, if the group persistently views you as essentially a marginal member, the concomitant lack of acceptance and inclusion leaves you feeling continually uncertain about your membership status. Typically this weakens ties to the group and ultimate leads to disidentification.

Extremism

One of the most far-reaching extensions and implications of uncertainty-identity theory is its analysis of extremism (Hogg, 2004, 2005). There is a substantial literature that documents a relationship between societal uncertainty and various forms of extremism, such as “totalist” groups (Baron et al., 2003), cults (Curtis and Curtis, 1993), genocide (Staub, 1989), terrorism (Moghaddam and Marsella, 2004), fascism (Billig, 1978), ultranationalism (Kosterman and Feshbach, 1989), blind patriotism (Staub, 1997), religious fundamentalism (Altemeyer, 2003; Rowatt and Franklin, 2004), authoritarianism (Doty et al., 1991), ideological thinking (Billig, 1982; Jost and Hunyady, 2002; Jost et al., 2003; Lambert et al., 1999), and fanaticism and being a “true believer” [p. 73 ↓] (Hoffer, 1951) – also see Hogg and Blaylock (in press). Uncertainty-identity theory can describe a process that generates extremism from uncertainty

Extremist groups have closed and carefully policed boundaries, uniform attitudes, values and membership, and inflexible customs. They are rigidly and hierarchically structured with a clearly delineated chain of legitimate influence and command, and substantial intolerance of internal dissent and criticism. Such groups are often ethnocentric, inward looking, and suspicious and disparaging of outsiders. They engage in relatively asocial and overly assertive actions that resemble collective narcissism (cf. Baumeister et al., 1996; Golec de Zavala et al., 2009): grandiosity, self-importance, envy, arrogance, haughtiness, entitlement, exploitativeness, excessive admiration, lack of empathy, fantasies of unlimited success, and feelings of special/unique/high status. In a similar vein, Kruglanski and colleagues describe a constellation of behaviors called “group-centrism,” which emerges

[w]hen people care a lot about sharing opinions with others in their group; when they endorse central authority that sets uniform norms and standards; when they suppress dissent, shun diversity, and show in-group favoritism; when they venerate their group's norms and traditions, and display fierce adherence to its views; when above all, they exhibit all these as a package.

Kruglanski et al., 2006: 84

These are “extreme” groups that, even if they have only some of the attributes described above, furnish members with an all-embracing, rigidly defined, exclusive, and highly prescriptive social identity and sense of self.

Not all uncertainty drives people into the arms of totalist groups. Such groups can be uncomfortably constraining as they are often authoritarian and dictate and control every aspect of one's life and identity. However, these groups may seem particularly attractive under conditions of extreme and enduring uncertainty; for example, widespread societal uncertainty caused by economic collapse, cultural disintegration, civil war, terrorism, and large-scale natural disasters; or more personal uncertainty caused by unemployment, bereavement, divorce, relocation, adolescence, and so forth.

Under these conditions totalist groups do an excellent job of reducing self-uncertainty. They are distinctive with rigid boundaries, often policed by the group, that unequivocally define who is in and who is out – there is no ambiguity or fuzziness about membership. The group's identity is clearly, unambiguously, and relatively simply defined, and often sharply polarized away from other groups. As a member you know exactly who you are and how you should behave and how others will behave. There is strong expectation of homogeneity and consensus that provides powerful social validation of one's identity and worldview, but also encourages a silo mentality in which dissenters and critics are suppressed and vilified. Such groups tend to be insular and inward looking; which provides a comfortably circumscribed world for members, but is also associated with marked ethnocentrism (Brewer and Campbell, 1976), accentuated mistrust and fear of outsiders (Stephan and Stephan, 1985), and a powerful tendency toward essentialism (Haslam et al., 1998) that renders self and social context subjectively stable and immutable.

Orthodoxy prevails (e.g., Deconchy, 1984). There is a single absolute standard of right and wrong in which attitudes, values, and behaviors are tightly woven together into ideological belief systems that are self-contained and explanatory (Larraín, 1979; Thompson, 1990); providing a firm and unassailable platform of certitude. The conjunction of moral absolutism, ideological orthodoxy and ethnocentrism is often a powerful basis for treating outgroup members as less than human – a process of dehumanization that can have terrible consequences (Haslam, 2006; Haslam et al., 2008).

Rigid ideological systems are particularly attractive in a postmodern world of moral [p. 74 ↓] and behavioral relativities and “limitless” choice. They resolve what Dunn (1998) has called the postmodern paradox: individual freedom of choice brings with it uncertainty about what to do and who to be and thus a desperate yearning for moral absolutes embedded in groups circumscribed by powerful ideologies. Religion has always provided all-embracing orthodoxies that not only address day-to-day uncertainties but also existential uncertainty – religious identification is such a powerful resolution of self-uncertainty that it can often mutate into zealotry associated, ironically, with intolerance (Hogg, Adelman et al., 2010; McGregor et al., 2008). For example, Lewis (2004), in his analysis of contemporary Islamic fundamentalism, argues that “in a time of intensifying strains, of faltering ideologies, jaded loyalties, and crumbling institutions, an ideology expressed in Islamic terms” (Lewis, 2004: 19) is particularly appealing.

We have already discussed, above, the important role of group leadership in reducing uncertainty, and how leaders can strategically use uncertainty to maintain their position of influence in the group. In extremist groups leadership becomes even more important. These groups are rigidly and consensually structured in terms of relative prototypicality of members, and of course prototype information is supremely important. Ideal conditions exist for prototypical leadership to prevail, and for such leaders to be extremely influential. Ultimately such leaders can become intoxicated by their power and feel isolated from the rank-and-file of the group – they can all too easily become autocratic despots (Hogg, 2007b). History is replete with examples: Adolf Hitler, Idi Amin, Saddam Hussein, Pol Pot, and so forth.

To some extent extremist groups are simply ultraentitative groups, and that is why uncertainty, particularly more extreme uncertainty, can make them attractive and why people identify very strongly with them. However, the construct of entitativity is primarily a perceptual construct that describes group structure – it does not speak to what a group does, the extent to which a group adopts a moderate or more radical course of action to protect or promote its identity and the welfare of its members. Extremist groups often have a powerful behavioral dimension focused on endorsement of and engagement in radical action.

This action component of a group's identity is likely to become more important to the extent that what the group stands for is self-relevant and viewed as under threat. When people feel their security, prosperity, and lifestyle are threatened, they yearn to identify strongly with a group that can get things done to remove or buffer the threat – a radical extremist group that has a forceful behavioral agenda. Against this background self-uncertainty not only strengthens identification with assertive radical groups, perhaps transforming members into fanatics, zealots, true believers, and ideologues, but also weakens identification with less assertive moderate groups. In this way identification with extreme groups may be a powerful force for social mobilization that transforms attitudes into action (e.g., Hogg and Smith, 2007; Klandermans, 1997; Stürmer and Simon, 2004).

Preliminary support for this idea comes from an experiment by Hogg, Meehan et al. (2010). In the context of a self-relevant threat, uncertainty strengthened students' identification with a radical campus action group and weakened identification with a more moderate group – identification also mediated intentions to engage in behavior on behalf of the group (also see Hogg, Farquharson, et al, 2010). Further support comes from four field studies conducted by Adelman et al. (submitted) in Israel. Palestinian Muslims and Israeli Jews, with stronger, more important, and more central national and religious identities, indicated greater support for violent action under high than low uncertainty.

[p. 75 ↓]

Relevance and Social Engagement

Uncertainty-identity theory is not merely an academic account of how feelings of self-uncertainty can motivate group identification. It has far-reaching relevance for understanding and engaging with important social issues. For example, as discussed above, it can help explain the enduring appeal of religious fundamentalism and ideological orthodoxy (Hogg, 2004, 2005; Hogg, Adelman et al., 2010), the conditions under which leadership can become autocratic or despotic (e.g., Hogg, 2007b), why particular groups may marginalize, suppress, and persecute dissenters or critics, and why groups develop an ethnocentric silo mentality that can dehumanize outgroups. It has also been used to explain why western adolescents may identify with extreme adolescent groups that prescribe dangerous behaviors that place their health at risk (Hogg, Siegel and Hohman, *in press*).

It may also be able to explain why specific individuals resort to terrorism: against a background of identity threat, perceived relative deprivation and uncertain times, elevated self-uncertainty may create a desperate yearning to belong and to do whatever it takes to promote protect and stabilize one's social identity, even engage in extreme violence against innocents if that is thought to be endorsed by the group's leadership. This behavior may also be seen as a way to gain validation of one's identity in the group and to be viewed as a bone fide core member – the oft-witnessed zealotry of neophytes and true believers.

Conclusion

In this chapter I have described uncertainty-identity theory (e.g., Hogg, 2000, 2007a) – a personal narrative of its origins and development; a description of its basic concepts, processes, and phenomena; a description of implications, extensions, and applications; and a short summary of its social relevance. As this is a book on theories, the emphasis has been conceptual, with empirical evidence and issues only briefly referenced as they have been overviewed elsewhere (e.g., Hogg, 2000, 2007a).

Uncertainty-identity theory proposes that self-uncertainty reduction is a key motivation for social identity processes and group and intergroup behaviors. It is a theory that attributes particular forms of group attachment, self-definition, and group structure to people's striving to reduce, via group identification, self-categorization and prototype-based depersonalization, feelings of uncertainty about and related to themselves. The core features of uncertainty-identity theory can be captured by three broad premises.

- *Premise 1.* People are motivated to reduce or avoid feelings of uncertainty about themselves, and about their perceptions, judgments, attitudes, and behaviors that relate to themselves, their interactions with other people, and their place in social context.
- *Premise 2.* Social categorization reduces or protects from uncertainty because it depersonalizes perception to conform to one's ingroup and outgroup prototypes, such that one "knows" how others will behave. Prototypes define and prescribe people's identities and therefore their perceptions, attitudes, feelings, and behaviors, and how they interact with and treat other people, including oneself. Social categorization of self, self-categorization, assigns one an identity with all its associated ingroup prototypical attributes. There is usually substantial agreement within a group on the ingroup prototype and on prototypes of relevant outgroups, further reducing uncertainty through consensual validation of one's behaviors and sense of self.
- *Premise 3.* Prototypes are better at resolving uncertainty to the extent that they are simple, clear, unambiguous, prescriptive, focused, and consensual, as well as coherently integrated, self-contained, and explanatory. These kinds of prototypes circumscribe clear identities and define or are associated with distinctive, well-structured groups that are high in entitativity. Under uncertainty people identify more strongly with high entitativity groups – they seek them out to join, they create them anew, or they transform existing groups to be more entitative.

These core features are the foundation for a number of elaborations and extensions; for example, relating to social influence and group norms, leadership and trust, and dissent, deviance, and minorities. Perhaps the most significant elaboration and extension is the theory's ability to help explain the emergence of social extremism. Where uncertainty

is extreme and enduring the motivation to reduce uncertainty and the quest for high entitativity groups and clear prototypes are strengthened. Under these circumstances people may identify passionately as true believers or zealots, seeking rigidly and hierarchically structured totalist groups with closed boundaries, homogenous and ideological belief structures, inflexible customs, and radical agendas – ethnocentric, insular and somewhat narcissistic groups that suppress dissent, are intolerant of outsiders, and engage in radical actions. These kinds of groups provide all-embracing identities that are powerful buffers against self-uncertainty.

Uncertainty is a pervasive part of life – we get excited and stimulated by it, we get frightened and oppressed by it, and we do what we can to reduce, control, or avoid it. We can never be truly certain so we are always more or less uncertain. In this chapter I have described a theory of how uncertainty may be related to why and how we identify with groups and to the particular types of groups that we identify with, suggesting that extreme uncertainty may encourage zealotry and totalism. In terms of social engagement, it goes without saying that these last are the bane of human existence – at best producing inefficient and oppressive groups; at worst, causing immeasurable human suffering.

References

Abrams, D. and Hogg, M.A. Comments on the motivational status of self-esteem in social identity and intergroup discrimination . European Journal of Social Psychology , 18, 317–334. (1988) <http://dx.doi.org/10.1002/ejsp.2420180403>

Abrams, D. and Hogg, M.A. Metatheory: Lessons from social identity research . Personality and Social Psychology Review , 8, 98–106. (2004) http://dx.doi.org/10.1207/s15327957pspr0802_2

Abrams, D. and Hogg, M.A. (2010) Social identity and self-categorization . In J.F. Dovidio, ed. , M. Hewstone, ed. , P. Glick, ed. and V.M. Esses (eds), The SAGE Handbook of Prejudice, Stereotyping, and Discrimination , pp. 179–193. London: Sage. <http://dx.doi.org/10.4135/9781446200919>

Adelman, J.R., Hogg, M.A. and Levin, S. (submitted) Support for political action as a function of religiousness and nationalism under uncertainty: A study of Muslims, Jews, Palestinians and Israelis in Israel . Manuscript submitted for publication.

Altemeyer, B. Why do religious fundamentalists tend to be prejudiced . International Journal for the Psychology of Religion , 13, 17–28. (2003) http://dx.doi.org/10.1207/S15327582IJPR1301_03

Arkin, R.M., ed. , Oleson, K.C., ed. and Carroll, P.J. (eds) (2010) Handbook of the Uncertain Self . New York: Psychology Press.

Baron, R.S., Crawley, K. and Paulina, D. (2003) Aberrations of power: Leadership in totalist groups . In D. van Knippenberg, ed. and M.A. Hogg (eds), Leadership and Power: Identity Processes in Groups and Organizations , pp. 169–183. London: Sage. <http://dx.doi.org/10.4135/9781446216170>

Bartlett, F.C. (1932) Remembering . Cambridge: Cambridge University Press.

Baumeister, R.F., Smart, L. and Boden, J.M. Relation of threatened egotism to violence and aggression: The dark side of high self-esteem . Psychological Review , 103, 5–33. (1996) <http://dx.doi.org/10.1037/0033-295X.103.1.5>

Berger, C.R. (1987) Communicating under uncertainty . In M.E. Roloff, ed. and G.R. Miller (eds), Interpersonal Processes: New Directions in Communication Research , pp. 39–62. Newbury Park, CA: Sage.

Billig, M. (1978) Fascists: A Social Psychological View of the National Front . London: Harcourt Brace Jovanovich.

Billig, M. (1982) Ideology and Social Psychology: Extremism, Moderation and Contradiction . London: Sage.

Blascovich, J. and Tomaka, J. (1996) The biopsychosocial model of arousal regulation . In M. Zanna (ed.), Advances in Experimental Social Psychology, Vol. 28 , 1–51. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960235-X>

Brewer, M.B. The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues*, 55, 429–444. (1999) <http://dx.doi.org/10.1111/0022-4537.00126>

Brewer, M.B. and Campbell, D.T. (1976) Ethnocentrism and Intergroup Attitudes: East African Evidence . New York: Sage.

Campbell, D.T. Common fate, similarity, and other indices of the status of aggregates of persons as social entities . *Behavioral Science* , 3, 14–25. (1958) <http://dx.doi.org/10.1002/bs.3830030103>

Castano, E., Yzerbyt, V.Y. and Bourguignon, D. We are one and I like it: The impact of ingroup entitativity on ingroup identification . *European Journal of Social Psychology* , 33, 735–754. (2003) <http://dx.doi.org/10.1002/ejsp.175>

Crano, W.D. and Seyranian, V. How minorities prevail: The context/comparison-lenienty contract model . *Journal of Social Issues* , 65, 335–363. (2009) <http://dx.doi.org/10.1111/j.1540-4560.2009.01603.x>

Curtis, J.M. and Curtis, M.J. Factors related to susceptibility and recruitment by cults . *Psychological Reports* , 73, 451–460. (1993) <http://dx.doi.org/10.2466/pr0.1993.73.2.451>

David, B. and Turner, J.C. (2001) Majority and minority influence: A single process self-categorization analysis . In C.K.W. De Dreu, ed. and N.K. De Vries (eds), *Group Consensus and Innovation* , pp. 91–121. Oxford: Blackwell.

Deconchy, J.P. (1984) Rationality and social control in orthodox systems . In H. Tajfel, (ed.), *The Social Dimension: European Developments in Social Psychology* , 2, 425–445. Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511759154.004>

Dewey, J. (1929/2005) *The Quest for Certainty: A Study of the Relation of Knowledge and Action* . Whitefish, MT: Kessinger Publishing.

Doty, R.M., Peterson, B.E. and Winter, D.G. Threat and authoritarianism in the United States, 1978–1987 . *Journal of Personality and Social Psychology* , 61, 629–640. (1991) <http://dx.doi.org/10.1037/0022-3514.61.4.629>

Dunn, R.G. (1998) Identity Crises: A Social Critique of Postmodernity . Minneapolis: University of Minnesota Press.

Festinger, L. A theory of social comparison processes . *Human Relations* , 7, 117–140. (1954) <http://dx.doi.org/10.1177/001872675400700202>

Fiske, S.T. and Taylor, S.E. (1991) Social Cognition , 2nd Edition. New York: McGraw-Hill.

Fromm, E. (1947) Man for Himself: An Inquiry into the Psychology of Ethics . New York: Rinehart.

Golec de Zavala, A., Cichocka, A., Eidelson, R. and Jayawickreme, N. Collective narcissism and its social consequences . *Journal of Personality and Social Psychology* , 97, 1074–1096. (2009) <http://dx.doi.org/10.1037/a0016904>

Grieve, P. and Hogg, M.A. Subjective uncertainty and intergroup discrimination in the minimal group situation . *Personality and Social Psychology Bulletin* , 25, 926–940. (1999) <http://dx.doi.org/10.1177/01461672992511002>

Hamilton, D.L. and Sherman, S.J. Perceiving persons and groups . *Psychological Review* , 103, 336–355. (1996) <http://dx.doi.org/10.1037/0033-295X.103.2.336>

Haslam, N. Dehumanization: An integrative review . *Personality and Social Psychology Review* , 10, 252–264. (2006) http://dx.doi.org/10.1207/s15327957pspr1003_4

Haslam, N., Loughnan, S. and Kashima, Y. Attributing and denying humanness to others . *European Review of Social Psychology* , 19, 55–85. (2008) <http://dx.doi.org/10.1080/10463280801981645>

Haslam, N., Rothschild, L. and Ernst, D. Essentialist beliefs about social categories . *British Journal of Social Psychology* , 39, 113–127. (1998) <http://dx.doi.org/10.1348/014466600164363>

Heine, S.J., Proulx, T. and Vohs, K.D. The meaning maintenance model: On the coherence of social motivations . *Personality and Social Psychology Review* , 10, 88–111. (2006) http://dx.doi.org/10.1207/s15327957pspr1002_1

Higgins, E.T. (1998) Promotion and prevention: Regulatory focus as a motivational principle . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 30, 1–46. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960381-0>

Hoffer, E. (1951) *The True Believer* . New York: Time.

Hofstede, G. (1980) *Culture's Consequences: International Differences in Work-related Values* . Beverly Hills, CA: Sage.

Hogg, M.A. Group cohesiveness: A critical review and some new directions . *European Review of Social Psychology* , 4, 85–111. (1993) <http://dx.doi.org/10.1080/14792779343000031>

Hogg, M.A. Subjective uncertainty reduction through self-categorization: A motivational theory of social identity processes . *European Review of Social Psychology* , 11, 223–255. (2000) <http://dx.doi.org/10.1080/14792772043000040>

Hogg, M.A. A social identity theory of leadership . *Personality and Social Psychology Review* , 5, 184–200. (2001) http://dx.doi.org/10.1207/S15327957PSPR0503_1

Hogg, M.A. (2004) Uncertainty and extremism: Identification with high entitativity groups under conditions of uncertainty . In V. Yzerbyt, ed. , C.M. Judd, ed. and O. Corneille (eds), *The Psychology of Group Perception: Perceived Variability, Entitativity, and Essentialism* , pp. 401–418. New York: Psychology Press.

Hogg, M.A. (2005) Uncertainty, social identity and ideology . In S.R. Thye, ed. and E.J. Lawler (eds), *Advances in Group Processes* , 22, 203–230. New York: Elsevier. <http://dx.doi.org/10.1016/S0882-6145%2805%2922008-8>

Hogg, M.A. (2006) Social identity theory . In P.J. Burke (ed.), *Contemporary Social Psychological Theories* , pp. 111–136. Palo Alto: Stanford University Press.

Hogg, M.A. (2007a) Uncertainty-identity theory . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 39, 69–126. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2806%2939002-8>

Hogg, M.A. (2007b) Organizational orthodoxy and corporate autocrats: Some nasty consequences of organizational identification in uncertain times . In C.A. Bartel, ed. , S. Blader, ed. and A. Wrzesniewski (eds), *Identity and the Modern Organization* , pp. 35–59. Mahwah, NJ: Erlbaum.

Hogg, M.A. (2008) Personality, individuality, and social identity . In F. Rhodewalt (ed.), *Personality and Social Behavior* , pp. 177–196. New York: Psychology Press.

Hogg, M.A. (2010a) Human groups, social categories, and collective self: Social identity and the management of self-uncertainty . In R.M. Arkin, ed. , K.C. Oleson, ed. , and P.J. Carroll (eds), *Handbook of the Uncertain Self* , pp. 401–420. New York: Psychology Press.

Hogg, M.A. (2010b) Influence and leadership . In S.T. Fiske, ed. , D.T. Gilbert, ed. and G. Lindzey (eds), *Handbook of Social Psychology* , 5th Edition, 2, 1166–1207. New York: Wiley.

Hogg, M.A. and Abrams, D. (1988) *Social Identifications: A Social Psychology of Intergroup Relations and Group Processes* . London: Routledge.

Hogg, M.A. and Abrams, D. (1993) Towards a single-process uncertainty-reduction model of social motivation in groups . In M.A. Hogg, ed. and D. Abrams, (eds), *Group Motivation: Social Psychological Perspectives* , pp. 173–190. Hemel Hempstead: Harvester Wheatsheaf/ New York: Prentice Hall.

Hogg, M.A., Adelman, J.R. and Blagg, R.D. Religion in the face of uncertainty: An uncertainty-identity theory account of religiousness . *Personality and Social Psychology Review* , 14, 72–83. (2010) <http://dx.doi.org/10.1177/1088868309349692>

Hogg, M.A., ed. and Blaylock, D.L. (eds) (in press) *Extremism and the Psychology of Uncertainty*. Boston: Wiley-Blackwell.

Hogg, M.A., Farquharson, J., Parsons, A. and Svensson, A. (2010) When being moderate is not the answer: Disidentification with moderate groups under uncertainty. Unpublished manuscript, Claremont Graduate University.

Hogg, M.A., Fielding, K.S. and Darley, J. (2005) Fringe dwellers: Processes of deviance and marginalization in groups. In D. Abrams, ed. , M.A. Hogg, ed. , and J.M. Marques (eds), *The Social Psychology of Inclusion and Exclusion* , pp. 191–210. New York: Psychology Press.

Hogg, M.A., Meehan, C. and Farquharson, J. The solace of radicalism: Self-uncertainty and group identification in the face of threat. *Journal of Experimental Social Psychology* , 46, 1061–1066. (2010) <http://dx.doi.org/10.1016/j.jesp.2010.05.005>

Hogg, M.A. and Mullin, B.-A. (1999) Joining groups to reduce uncertainty: Subjective uncertainty reduction and group identification. In D. Abrams, ed. and M.A. Hogg (eds), *Social Identity and Social Cognition* , pp. 249–279. Oxford: Blackwell.

Hogg, M.A. and Reid, S.A. Social identity, self-categorization, and the communication of group norms. *Communication Theory* , 16, 7–30. (2006) <http://dx.doi.org/10.1111/j.1468-2885.2006.00003.x>

Hogg, M.A., Sherman, D.K., Dierselhuis, J., Maitner, A.T. and Moffitt, G. Uncertainty, entitativity, and group identification. *Journal of Experimental Social Psychology* , 43, 135–142. (2007) <http://dx.doi.org/10.1016/j.jesp.2005.12.008>

Hogg, M.A., Siegel, J.T. and Hohman, Z.P. (in press) Groups can jeopardize your health: Identifying with un-healthy groups to reduce self-uncertainty. *Self and Identity* .

Hogg, M.A. and Smith, J.R. Attitudes in social context: A social identity perspective. *European Review of Social Psychology* , 18, 89–131. (2007) <http://dx.doi.org/10.1080/10463280701592070>

Hogg, M.A. and Svensson, A. (2010) Uncertainty, self-esteem and group identification . Unpublished manuscript, Claremont Graduate University.

Hogg, M.A. and van Knippenberg, D. (2003) Social identity and leadership processes in groups . In M.P. Zanna (ed.), Advances in Experimental Social Psychology, Vol. 35 , 1–52. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2803%2901001-3>

Hohman, Z.P., Hogg, M.A. and Bligh, M.C. Identity and intergroup leadership: Asymmetrical political and national identification in response to uncertainty . Self and Identity , 9, 113–128. (2010) <http://dx.doi.org/10.1080/15298860802605937>

Hornsey, M.J. Why being right is not enough: Predicting defensiveness in the face of group criticism . European Review of Social Psychology , 16, 301–334. (2005) <http://dx.doi.org/10.1080/10463280500436040>

Jetten, J., Hogg, M.A. and Mullin, B.-A. Ingroup variability and motivation to reduce subjective uncertainty . Group Dynamics: Theory Research, and Practice , 4, 184–198. (2000) <http://dx.doi.org/10.1037/1089-2699.4.2.184>

Jost, J.T., Glaser, J., Kruglanski, A.W. and Sulloway, F.J. Political conservatism as motivated social cognition . Psychological Bulletin , 129, 339–375. (2003) <http://dx.doi.org/10.1037/0033-2909.129.3.339>

Jost, J.T. and Hunyady, O. The psychology of system justification and the palliative function of ideology . European Review of Social Psychology , 13, 111–153. (2002) <http://dx.doi.org/10.1080/10463280240000046>

Kahneman, D., ed. , Slovic, P., ed. and Tversky, A. (eds) (1982) Judgment Under Uncertainty: Heuristics and Biases . New York: Cambridge University Press.

Klandermans, B. (1997) The Social Psychology of Protest . Oxford: Blackwell.

Koffka, K. (1935) Principles of Gestalt Psychology . New York: Harcourt, Brace and Co.

Kosterman, R. and Feshbach, S. Towards a measure of patriotic and nationalistic attitudes . *Political Psychology* , 10, 257–274. (1989) <http://dx.doi.org/10.2307/3791647>

Kruglanski, A.W., Pierro, A., Mannetti, L. and De Grada, E. Groups as epistemic providers: Need for closure and the unfolding of group-centrism . *Psychological Review* , 113, 84–100. (2006) <http://dx.doi.org/10.1037/0033-295X.113.1.84>

Kruglanski, A.W. and Webster, D.M. Motivated closing of the mind: ‘Seizing’ and ‘freezing’ . *Psychological Review* , 103, 263–283. (1996) <http://dx.doi.org/10.1037/0033-295X.103.2.263>

Kuhn, T. (1962) *The Structure of Scientific Revolutions* . Chicago, IL: University of Chicago Press.

Lambert, A.J., Burroughs, T. and Nguyen, T. Perceptions of risk and the buffering hypothesis: The role of just world beliefs and right-wing authoritarianism . *Personality and Social Psychology Bulletin* , 25, 643–656. (1999) <http://dx.doi.org/10.1177/0146167299025006001>

Larraín, J. (1979) *The Concept of Ideology* . London: Hutchinson.

Lewis, B. (2004) *The Crisis of Islam: Holy War and Unholy Terror* . London: Phoenix.

Lickel, B., Hamilton, D.L., Wieczorkowska, G., Lewis, A., Sherman, S.J. and Uhles, A.N. Varieties of groups and the perception of group entitativity . *Journal of Personality and Social Psychology* , 78, 223–246. (2000) <http://dx.doi.org/10.1037/0022-3514.78.2.223>

Lind, E.A. and Tyler, T.R. (1988) *The Social Psychology of Procedural Justice* . New York: Plenum Press.

Marques, J.M., Abrams, D. and Serôdio, R. Being better by being right: Subjective group dynamics and derogation of in-group deviants when generic norms are undermined . *Journal of Personality and Social Psychology* , 81, 436–447. (2001) <http://dx.doi.org/10.1037/0022-3514.81.3.436>

Marques, J.M. and Paez, D. The 'black sheep effect': Social categorization, rejection of ingroup deviates and perception of group variability . European Review of Social Psychology , 5, 37–68. (1994) <http://dx.doi.org/10.1080/14792779543000011>

Marris, P. (1996) The Politics of Uncertainty: Attachment in Private and Public Life . London: Routledge.

Martin, R. and Hewstone, M. (2008) Majority versus minority influence, message processing and attitude change: The source-context-elaboration model . In M.P. Zanna (ed.), Advances in Experimental Social Psychology , 40, 237–326. San Diego: Elsevier. <http://dx.doi.org/10.1016/S0065-2601%2807%2900005-6>

Maslow, A.H. (1987) Motivation and Personality , 3rd Edition. New York: Harper Collins.

McGarty, C., Turner, J.C., Oakes, P.J. and Haslam, S.A. The creation of uncertainty in the influence process: The roles of stimulus information and disagreement with similar others . European Journal of Social Psychology , 23, 17–38. (1993) <http://dx.doi.org/10.1002/ejsp.2420230103>

McGregor, I., Haji, R., Nash, K.A. and Teper, R. Religious zeal and the uncertain self . Basic and Applied Social Psychology , 30, 183–188. (2008) <http://dx.doi.org/10.1080/01973530802209251>

McGregor, I. and Marigold, D.C. Defensive zeal and the uncertain self: What makes you so sure? Journal of Personality and Social Psychology , 85, 838–852. (2003) <http://dx.doi.org/10.1037/0022-3514.85.5.838>

McGregor, I., Zanna, M.P., Holmes, J.G. and Spencer, S.J. Compensatory conviction in the face of personal uncertainty: Going to extremes and being oneself . Journal of Personality and Social Psychology , 80, 472–488. (2001) <http://dx.doi.org/10.1037/0022-3514.80.3.472>

Moghaddam, F.M., ed. and Marsella, A.J. (eds) (2004) Understanding Terrorism: Psychosocial Roots, Consequences, and Interventions . Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10621-000>

Moscovici, S. (1980) Toward a theory of conversion behavior . In L. Berkowitz (ed.), Advances in Experimental Social Psychology , 13, 202–239. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960133-1>

Mugny, G. and Pérez, J. (1991) The Social Psychology of Minority Influence . Cambridge: Cambridge University Press.

Mullin, B.-A. and Hogg, M.A. Dimensions of subjective uncertainty in social identification and minimal intergroup discrimination . British Journal of Social Psychology , 37, 345–365. (1998) <http://dx.doi.org/10.1111/j.2044-8309.1998.tb01176.x>

Oakes, P.J. (1987) The salience of social categories . In J.C. Turner, ed. , M.A. Hogg, ed. , P.J. Oakes, ed. , S.D. Reicher, ed. and M.S. Wetherell, (eds), Rediscovering the Social Group: A Self-categorization Theory , pp. 117–141. Oxford: Blackwell.

Pickett, C.L. and Brewer, M.B. Assimilation and differentiation needs as motivational determinants of perceived ingroup and outgroup homogeneity . Journal of Experimental Social Psychology , 37, 341–348. (2001) <http://dx.doi.org/10.1006/jesp.2000.1469>

Pickett, C.L., Silver, M.D. and Brewer, M.B. The impact of assimilation and differentiation needs on perceived group importance and judgments of ingroup size . Personality and Social Psychology Bulletin , 28, 546–558. (2002) <http://dx.doi.org/10.1177/0146167202287011>

Pollock, H.N. (2003) Uncertain Science ... Uncertain World . Cambridge: Cambridge University Press.

Reid, S.A. and Hogg, M.A. Uncertainty reduction, self-enhancement, and ingroup identification . Personality and Social Psychology Bulletin , 31, 804–817. (2005) <http://dx.doi.org/10.1177/0146167204271708>

Rowatt, W.C. and Franklin, L.M. Christian orthodoxy, religious fundamentalism, and right-wing authoritarianism as predictors of implicit racial prejudice . International Journal for the Psychology of Religion , 14, 125–138. (2004) http://dx.doi.org/10.1207/s15327582ijpr1402_4

Saks, A.M. and Ashforth, B.E. Organizational socialization: Making sense of the past and present as a prologue for the future . *Journal of Vocational Behavior* , 51, 234–279. (1997) <http://dx.doi.org/10.1006/jvbe.1997.1614>

Sani, F. and Reicher, S.D. When consensus fails: An analysis of the schism within the Italian Communist Party (1991) . *European Journal of Social Psychology* , 28, 623–45. (1998) <http://dx.doi.org/10.1002/%28SICI%291099-0992%28199807/08%2928:4%3C623::AID-EJSP885%3E3.0.CO;2-G>

Schwartz, S.H. (1992) Universals in the content and structure of values: Theoretical advances and empirical tests in 20 cultures . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 25, 1–65. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960281-6>

Sherman, D.K., Hogg, M.A. and Maitner, A.T. Perceived polarization: Reconciling ingroup and intergroup perceptions under uncertainty . *Group Processes and Intergroup Relations* , 12, 95–109. (2009) <http://dx.doi.org/10.1177/1368430208098779>

Sorrentino, R.M. and Roney, C.J.R. (1999) *The Uncertain Mind: Individual Differences in Facing the Unknown* . Philadelphia, PA: Psychology Press.

Staub, E. (1989) *The Roots of Evil: The Psychological and Cultural Origins of Genocide and Other Forms of Group Violence* . New York: Cambridge University Press.

Staub, E. (1997) Blind versus constructive patriotism: Moving from embeddedness in the group to critical loyalty and action . In D. Bar-Tal, ed. and E. Staub (eds), *Patriotism: In the Lives of Individuals and Nations* , pp. 213–228. Chicago: Nelson-Hall.

Stephan, W.G. and Stephan, C.W. Intergroup anxiety . *Journal of Social Issues* , 41, 157–75. (1985) <http://dx.doi.org/10.1111/j.1540-4560.1985.tb01134.x>

Stürmer, S. and Simon, B. Collective action: Towards a dual-pathway model . *European Review of Social Psychology* , 15, 59–99. (2004)

Suls, J., ed. and Wheeler, L. (eds) (2000) *Handbook of Social Comparison: Theory and Research* . New York: Kluwer/Plenum. <http://dx.doi.org/10.1007/978-1-4615-4237-7>

Swann, W.B.Jr., Rentfrow, P.J. and Guinn, J.S. (2003) Self-verification: The search for coherence . In M.R. Leary, ed. and J.P. Tangney (eds), *Handbook of Self and Identity* , pp. 367–383. New York: Guilford Press.

Tajfel, H. Quantitative judgement in social perception . *British Journal of Psychology* , 50, 16–29. (1959) <http://dx.doi.org/10.1111/j.2044-8295.1959.tb00677.x>

Tajfel, H. and Turner, J.C. (1979) An integrative theory of intergroup conflict . In W.G. Austin, ed. and S. Worchel (eds), *The Social Psychology of Inter group Relations* , pp. 33–47. Monterey, CA: Brooks/Cole.

Thompson, J.B. (1990) Ideology and Modern Culture: Critical Social Theory in the Era of Mass Communication . Stanford, CA: Stanford University Press.

Turner, J.C. (1991) Social Influence . Milton Keynes: Open University Press.

Turner, J.C. (1999) Some current issues in research on social identity and self-categorization theories . In N. Ellemers, ed. , R. Spears, ed. and B. Doosje (eds), *Social Identity* , pp. 6–34. Oxford: Blackwell.

Turner, J.C., Hogg, M.A., Oakes, P.J., Reicher, S.D. and Wetherell, M.S. (1987) Rediscovering the Social Group: A Self-categorization Theory . Oxford: Blackwell.

Turner, J.C., Oakes, P.J., Haslam, S.A. and McGarty, C.A. Self and collective: Cognition and social context . *Personality and Social Psychology Bulletin* , 20, 454–463. (1994) <http://dx.doi.org/10.1177/0146167294205002>

Tyler, T.R. The psychology of legitimacy: A relational perspective on voluntary deference to authorities . *Personality and Social Psychology Review* , 1, 323–345. (1997) http://dx.doi.org/10.1207/s15327957pspr0104_4

Tyler, T.R. and Lind, E.A. (1992) A relational model of authority in groups . In M.P. Zanna (ed), *Advances in Experimental Social Psychology* , 25, 115–191. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960283-X>

Van den Bos, K. Making sense of life: The existential self trying to deal with personal uncertainty . Psychological Inquiry , 20, 197–217. (2009) <http://dx.doi.org/10.1080/10478400903333411>

Van den Bos, K., Poortvliet, P.M., Maas, M., Miedema, J. and Van den Ham, E.-J. An enquiry concerning the principles of cultural norms and values: The impact of uncertainty and mortality salience on reactions to violations and bolstering of cultural worldviews . Journal of Experimental Social Psychology , 41, 91–113. (2005) <http://dx.doi.org/10.1016/j.jesp.2004.06.001>

Van Vugt, M. and Hart, C.M. Social identity as social glue: The origins of group loyalty . Journal of Personality and Social Psychology , 86, 585–598. (2004) <http://dx.doi.org/10.1037/0022-3514.86.4.585>

Yzerbyt, V., Castano, E., Leyens, J.-P. and Paladino, M.-P. The primacy of the ingroup: The interplay of entitativity and identification . European Review of Social Psychology , 11, 257–295. (2000) <http://dx.doi.org/10.1080/14792772043000059>

<http://dx.doi.org/10.4135/9781446249222.n29>

Handbook of Theories of Social Psychology

Optimal Distinctiveness Theory: Its History and Development

Contributors: Marilynn B. Brewer

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Optimal Distinctiveness Theory: Its History and Development"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n30>

Print pages: 81-99

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n30>

[p. 81 ↓]

Chapter 30: Optimal Distinctiveness Theory: Its History and Development

Marilynn B.Brewer

Abstract

Optimal distinctiveness theory is a model of the motivations underlying attachment and identification with social groups. The theory posits that humans are characterized by two opposing needs that govern the relationship between the self and membership in social groups. The first is a need for assimilation and inclusion, a desire for belonging that motivates immersion in social groups. The second is a need for differentiation from others that operates in opposition to the need for immersion. As group membership becomes more and more inclusive, the need for inclusion is satisfied but the need for differentiation is activated; conversely, as inclusiveness decreases, the differentiation need is reduced but the need for assimilation is activated. According to the model, the two opposing motives produce an emergent characteristic – the capacity for social identification with distinctive groups that satisfy both needs simultaneously. The theory derived from a general perspective on the evolution of human sociality that recognizes that humans are adapted for group living and that the structural requirements for group cohesion and coordination have shaped social motivational systems at the individual level.

Introduction

When it comes to the intellectual history of my entire research career, all roads lead to Donald Campbell. When I entered the doctoral program in the Department of Psychology at Northwestern University in 1964, I was nominally admitted to study

social psychology. But the social psychology “program” at that time consisted solely of Don Campbell, who was unconvinced that disciplinary boundaries or area labels should constrain intellectual efforts or one’s scientific agenda (cf. Campbell, 1969). An intellectual giant, Don Campbell tackled big questions of epistemology, human evolution, and the sociology of science without regard for arbitrary distinctions between philosophy, biology, or the social sciences, and he encouraged his students to do the same. For me, as a recent graduate of a small liberal arts college, working with Don Campbell was a heady experience to say the least.

[p. 82 ↓]

While I was in graduate school, evolutionary biologist George Williams's influential book *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought* (1966) was published, followed a few years later by Edward O. Wilson's *Sociobiology* (1975). These two works became the backdrop for extended discussion and debate between Don and me – a debate that continued until his death in 1996. Convinced by Williams and Wilson that humans, like other organisms, are genetically selfish, Don believed that we had to look to the evolution of social institutions and powerful cultural and religious traditions to understand the social achievements of human beings. These ideas culminated in the text of his presidential address to the American Psychological Association in 1975, where he argued that there is an inherent conflict between the forces of biological evolution (selecting for individual self-interest) and those of social evolution (providing external constraints on selfishness in the interests of group survival). The implications, he suggested, were that psychologists and other social scientists should be wary of challenging moral traditions that have evolved to hold human selfishness in check (Campbell, 1975).

It was on these points that Don and I had our most interesting and challenging disagreements. I just could not accept the idea that the extent of sociality and sustained group living that characterizes human beings could have been maintained solely by external constraints embodied in social institutions, traditions, and practices selected at the group level and operating in opposition to biological selection. (As in all domains, Don nurtured the debate, encouraging me to develop and argue my own position, even though he rarely altered his own.) I expressed my disagreement in print in a short comment on his presidential address (Brewer, 1976; see also Brewer, 1989) where

I argued that the profound ambivalence between personal self-gratification and self-sacrifice for collective welfare is not a conflict between internal biological motives and external social constraints but rather an internal biological dualism that reflects human evolutionary history as a social species. My point was that human beings are neither inherently purely selfish nor purely altruistic but instead characterized by a kind of functional antagonism between self-interested and group-interested behavior. Because of the resultant variability in motives underlying human social behavior, I suggested, the distribution of human sociality might best be depicted in terms of a “golden standard deviation” rather than a “golden mean.”

Basically, Don and I did not disagree that there is a profound conflict between individual-level self-interest that drives inter-individual competition within social groups on the one hand, and the collective level interests that require cooperation and coordination transcending individual self-interest on the other. Where we disagreed was whether the locus of the conflict between selfishness and social cooperation lies in opposing forces of biological and social evolution or in an inherent dualism within our biological nature.

Although Don and I argued these issues sporadically over the years, the concepts of ambivalent sociality and opposing motives were latent influences that shaped some of my interest in social identity and group behavior but were not explicitly developed into more formal theory until 15 years later. It was the occasion of preparing my own presidential address for the Society of Personality and Social Psychology in 1990 that prompted me to formalize the idea of opposing motives in the form of optimal distinctiveness theory (Brewer, 1991). By that time I had become convinced that the product of the tension between human selfishness and human sociality was our capacity for intense identification with nonkin groups and its motivational underpinnings. Reaching that conclusion reflected the convergence of three different lines of research and theory that I had been exposed to in the ensuing 15 years.

[p. 83 ↓]

The Intellectual Ancestry of Optimal Distinctiveness Theory

Social Identity and Ethnocentrism

The first major influence that shaped my future thinking was my exposure to the study of ethnocentrism and ingroup identity during graduate school. At the same time that Don Campbell and I were initiating our debate about the nature of human sociality, we were also involved in an ambitious interdisciplinary project in collaboration with anthropologist Robert LeVine (then at the University of Chicago) to test cross-culturally the universality of ethnocentrism in human societies (see LeVine and Campbell, 1972). The term “ethnocentrism” was coined by William Graham Sumner in his book *Folkways* (1906). The concept was driven by the observation that human social arrangements are universally characterized by differentiation into ingroups and outgroups – the we–they distinctions that demarcate boundaries of loyalty and cooperation among individuals. Attitudes and values are shaped by this ingroup–outgroup distinction in that individuals view all others from the perspective of the ingroup. In Sumner’s words, ethnocentrism is

the view of things in which one's own group is the center of everything, and all others are scaled and rated with reference to it ... Each group nourishes its own pride and vanity, boasts itself superior, exalts its own divinities, and looks with contempt on outsiders. Each group thinks its own folkways the only right ones ... (E)thnocentrism leads a people to exaggerate and intensify everything in their own folkways which is peculiar and which differentiates them from others.

Sumner, 1906: 12–13

The Cross-Cultural Study of Ethnocentrism (CCSE) project (funded by a grant from the Carnegie Foundation) introduced a novel method of data collection designed to blend ethnographic case study and structured interview techniques. Experienced ethnographers in field sites in Africa, New Guinea, North America, and Asia were

commissioned to use their best local informants to obtain information on precolonial ingroup organization and intergroup attitudes, using a structured, open-ended interview format. Back in Evanston, Illinois, I took on the position of graduate research associate for the project, responsible for processing, organizing, and archiving the fieldnotes from each of the project sites as they were submitted by the ethnographers. That experience exposed me to the rich detail of ethnographic accounts of social behavior and provided exotic examples of customs, practices, and beliefs that reveal the enormous range of ways in which groups manage both intragroup and intergroup relationships. It also established in me a fascination with the study of group identity and intergroup attitudes that determined my research career path in social psychology from that point on.

The CCSE project did provide evidence relevant to Sumner's original hypotheses about the nature of ethnocentrism and human societies. Both qualitative and quantitative analyses of the coded interviews and survey responses from respondents in far-flung locations confirmed the robustness of the tendency to differentiate the social environment in terms of ingroup–outgroup distinctions and to value ingroup characteristics over those of other groups (Brewer, 1981, 1986; Brewer and Campbell, 1976). Importantly, the level of ingroup cohesion and loyalty did not appear to be correlated with degree of negativity of attitudes toward outgroups. Our interviews with representatives of non-Western societies revealed a wide range of attitudes toward recognized outgroups, from respect and mutual admiration to relative indifference to outright hostility. As one of our informants put it “[W]e have our ways and they have their ways,” and preference for the ingroup ways did not necessarily require intolerance of the outgroup. Thus, it was the experience gained from the CCSE project that first convinced me that ingroup preference and outgroup prejudice are two different constructs, with different origins and different consequences for intergroup behavior (Brewer, 1999, 2001). [p. 84 ↓] Contrary to Sumner's original analysis, I concluded that ingroup formation and attachment had their origins in factors other than intergroup conflict.

Meanwhile back in the laboratory, approximately simultaneous with the data collection phase of the CCSE project, Henri Tajfel's social psychology research group in Bristol, England, was developing a very different paradigm for studying ingroup bias and intergroup discrimination. Experiments with the so-called “minimal intergroup situation” (Tajfel, 1970; Tajfel et al., 1971) provided a powerful demonstration that

merely classifying individuals into arbitrary distinct social categories was sufficient to produce ingroup–outgroup discrimination and ingroup favoritism, even in the absence of any interactions with fellow group members or any history of competition or conflict between the groups. Additional experimental research demonstrated just how powerfully mere social categorization can influence thinking, feeling, and behaving toward ingroup members and the ingroup as a whole (Brewer, 1979).

Remarkably, results of the cross-cultural field research and these laboratory studies converged in confirming the power of we–they distinctions to produce differential evaluation, liking, and treatment of other persons depending on whether they are identified as members of the ingroup category or not. The laboratory experiments with the minimal intergroup situation demonstrated that ethnocentric loyalty and bias clearly do not depend on kinship or an extensive history of interpersonal relationships among group members, but can apparently be engaged readily by symbolic manipulations that imply shared attributes or common fate. Further, experiments with the minimal intergroup situation also provided evidence consistent with our CCSE data, that ingroup favoritism is prior to, and not necessarily associated with, outgroup negativity or hostility. What appears to be essential for ingroup attachment is that there be a basis for distinctive identification of who is “us” and who is “them” – a rule of exclusion as well as inclusion. The critical point drawn from the early experiments with the minimal group paradigm and ingroup favoritism was the evidence that individuals are willing to benefit fellow ingroup members even *in the absence of any direct self-interest or personal gain*.

Accounting for the results of the initial minimal group experiments and subsequent research on ingroup favoritism led to the development of social identity theory (Tajfel, 1981; Tajfel and Turner, 1979) which rests on the assumption that identification and emotional attachment to a social group redefines one's identity from the personal to the group level. Through the processes of self-categorization and group identification, an individual's sense of self and self-interest become inextricably tied to group interests and group welfare. In effect, social identity is a *transformation* of the self that redefines the meaning of self-interest (Brewer, 1991).

I had the opportunity to work with Henri Tajfel and his research group at the University of Bristol for a brief period in 1980, while the theory of social identity was being developed and tested. The remarkable convergence between the qualitative data from

our ethnographic field studies, results from survey studies, and findings from laboratory experiments on ingroup bias further convinced me that group identification is an inherent feature of human psychology that serves to regulate and maintain the essential relationship between individuals and their social groups. Social identity provides the constraint on human selfishness that makes cooperation and group existence possible. Understanding the nature of social identity and the motivations that drive and sustain individuals' attachment to their social groups, then, seemed to me a central task of social psychology.

Social Dilemmas and Collective Decision Making

The second major influence on my theoretical development came from an interest in the study of social dilemmas early in my [p. 85 ↓] academic career. A fundamental feature of group identification is the premise that when a social identity is activated, group interests and welfare supersede individual self-interest. In much of social life, individual self-interest and group interests coincide, so that cooperation and interdependence serve group goals and satisfy individual needs at the same time. If I desire the benefits of winning in a team sport competition, for instance, then cooperating with my fellow team members is clearly the best way for all of us to meet our individual and collective goals. But individual goals and collective interests do not always coincide so perfectly. If my individual interests are enhanced by being the one member of my team that scores the most points, but my team's chances of winning depend on my providing other team members the opportunity to score, working for my personal goal and achieving the group goal are not completely compatible. Whatever the long-term benefits of group living and cooperation may be, they often require mechanisms for overriding individual self-interests in the short-run. When individual self-interest and collective interests are placed in opposition, the innate ambivalences in human nature are revealed.

"Social dilemmas" constitute a special set of interdependence problems in which individual and collective interests are at odds. The dilemma arises whenever individuals acting in their own rational self-interest would engage in behaviors that cumulatively disadvantage everyone. A seminal article by Garret Hardin that appeared in *Science* in

1968 sparked interest in the study of individual decision making in the context of social dilemmas among behavioral economists, political scientists, and sociologists, as well as social psychologists. In his article, Hardin (1968) analyzed the parable of "the tragedy of the commons." The parable describes a situation in which a number of herdsmen graze their herds on a common pasture. Each individual herdsman is aware that it is to his benefit to increase the size of his herd because each animal represents profit to himself, while the cost of grazing the animal is shared by all the herdsmen. Responding to this incentive structure, each herdsman rationally decides to increase his herd size and, as a result, the commons deteriorates, the carrying capacity of the commons is exceeded, and ultimately leads to the collapse of the commons and the destruction of all of the herds that grazed on it.

Hardin's parable represents a form of social interdependence in which the collective consequence of reasonable self-interested individual choices is disaster. In the modern world, social dilemmas include problems of maintaining scarce collective resources such as water and rainforests, preserving public goods such as parks and public television, and preventing pollution and destruction of the environment. The self-interests of each individual are best served by taking advantage of the benefits of collective resources without contributing to their maintenance, but the cumulative effect of such self-interested actions would be that everyone pays the cost of resource depletion and environmental damage. To the extent that social life is characterized by these types of interdependencies, some mechanisms for balancing individual interests and collective welfare must be achieved.

I was introduced to the study of social dilemmas when I moved to the University of California at Santa Barbara in 1973 and had the opportunity to work with Charles McClintock and David Messick, whose research focused on social exchange and interdependence. For me, the structure of the *n*-person commons dilemma seemed the perfect forum for observing individual behavior when faced with a conflict between personal and collective interests. Together with a team of graduate students, Dave Messick and I developed a laboratory analogue of resource dilemmas (Parker et al., 1983), and Rod Kramer and I set out to explore the role of social identity in individual decision making in this resource dilemma context (Brewer and Kramer, 1986; Kramer and Brewer, 1984).

[p. 86 ↓]

Results from our own experiments and others indicate that in these choice situations, individuals do not behave consistently selfishly or unselfishly; a great deal depends on the group context in which the decision is made. When a collective social identification is not available, individuals tend to respond to the depletion of a collective resource by increasing their own resource use, at the cost of long-term availability. However, when a symbolic collective identity has been made salient, individuals respond to a resource crisis by dramatically reducing their own resource use (Brewer and Kramer, 1986; Brewer and Schneider, 1990). Further, when a public goods decision is preceded by even a brief period of group discussion, the rate of cooperative choice (when decisions are made individually and anonymously) is almost 100 percent (Caporael et al., 1989). This level of cooperative responding suggests that, under appropriate conditions, group welfare is just as "natural" as self-gratification as a rule for individual decision making. Situational cues, social identity salience, and behavior of others determine which predisposition will dominate on any particular occasion.

Our initial experiments demonstrated that self-sacrificial cooperative behavior in the interest of collective welfare is significantly increased when a salient social identity is shared by the interdependent group. Further experiments set out to determine the psychological mechanisms underlying this cooperative behavior. Willingness to contribute to collective welfare is determined in part by individuals' own social motives and in part by their expectations of what others will do in the dilemma situation. Shared ingroup identity influences both of these factors.

Self-sacrifice on behalf of the collective is wasted unless one has some trust that other members of the collective will also do their share. One function of ingroup formation and ingroup favoritism is providing a solution to the dilemma of social cooperation and trust (Brewer, 1986). Ordinarily, interpersonal trust depends on personal knowledge of other participants, such as a history of interpersonal exchange and future cooperation on an individualized basis. On the other hand, cooperation that is contingent on common membership in a bounded social group bypasses the need for such personalized knowledge or the costs of negotiating reciprocity with individual others. Shared ingroup membership affords a kind of "depersonalized trust" based on group membership (social identity) alone. All that is required for group-based trust and cooperation is (a)

the mutual knowledge that oneself and other share a common ingroup membership, and (b) the expectation that the other(s) will act in terms of that group membership in dealings with a fellow group member (and vice versa). In effect, one's own and other's behavior is perceived to be constrained by the requirements of group membership and the desire to retain one's status as an accepted group member. Ingroup trust is the expectation that others will cooperate with me *because we are members of the same group* (Foddy et al., 2008; Kramer and Wei, 1999; Tanis and Postmes, 2005).

Many social dilemmas (e.g., resource dilemmas, public goods contribution dilemmas) involve a decision whether or not to cooperate with the group as whole when one's own cooperative choice does not directly influence the cooperation of others. Under these circumstances, expecting that others will behave cooperatively (i.e., contribute to the public good or restrain consumption of a shared resource) reduces the fear that one's own cooperation will be wasted (i.e., the "sucker's payoff"). However, it does not eliminate the self-interested benefit of noncooperation. If everyone else can be expected to cooperate, then noncooperation takes advantage of the others' contributions to the group welfare and maximizes personal outcomes. Thus, expectations of others' intentions to cooperate are not of themselves sufficient to generate cooperative behavior. Ingroup trust can be exploited, particularly under conditions of anonymity and diffusion of responsibility. Group-based [p. 87 ↓] depersonalized trust translates to cooperative behavior only if the individual's own behavior is constrained by the same group norms that underlie his/her expectations of the others' behavior.

The psychological process of group identification, as elaborated in social identity theory, provides a basis for intragroup cooperation that does not necessarily rely on interpersonal trust in fellow group members. When individuals attach their sense of self to their group membership, they see themselves as interchangeable components of a larger social unit (Turner et al., 1987). The consequence of such social identification is not only affective attachment to the group as a whole, but also a shift of motives and values from self-interest to group interest and concern for the welfare of fellow group members. As a result of this redefinition of the self, pursuing the group's interest becomes a direct and natural expression of self-interest, that is, collective and personal interest are interchangeable. When the definition of self changes, the meaning of self-interest and self-serving motivations also changes accordingly. Group identity involves

a transformation of goals from the personal to the collective level (De Cremer and Van Vugt, 1999; Kramer and Brewer, 1986).

Goal transformation provides a basis for ingroup cooperation that does not depend directly on expectations that others in the group will reciprocate cooperation. When social identification is strong, then contributing to the group welfare is an end in itself, independent of what benefits ultimately accrue to the self. This is particularly evident when a group as a whole is failing to maintain a shared resource or public good, an indication that others in the group are not contributing sufficiently to group welfare. In the absence of strong, shared identity, indications that others are failing to cooperate is a cue to self-interested behavior that undermines any motive to cooperate. When group identification is strong, however, participants interpret negative group feedback as a signal that their group is in need and as such they should try harder at achieving their group goals. Consistent with the goal-transformation hypothesis, strong group identifiers exhibit a genuine concern for the group's welfare, and negative group feedback is interpreted as a threat to the group's welfare and a signal that behavioral changes are required, motivating them to cooperate more (Brewer and Schneider, 1990; DeCremer and van Dijk, 2002).

Years of research on how people behave in social dilemma situations both in the laboratory and in real life reinforced my conviction that there are internal psychological constraints on self-interest that are activated by group identification. What remained to be understood was the proximal mechanisms underlying the tie between individuals and their social groups.

Levels of Analysis and Downward Causation

A third influence on the development of optimal distinctiveness theory came from philosophy of science and evolutionary theory, as I was initially introduced to these disciplines by Don Campbell and later through my longstanding collaboration with Linnda Caporael (cf. Brewer and Caporael, 1990, 2006).

One of the problems with accounting for the evolution of self-sacrificial sociality is that the reproductive fitness value of such behaviors could not be modeled at the individual level and instead seemed to require some form of “group selection” mechanism. In its earliest form, group selection was proposed to explain why the size of populations remained within the carrying capacity of their environments (Wynne-Edwards, 1962). Presumably, some members of the population would sacrifice their own reproduction to benefit the group. But by Darwinian logic, genes that caused individuals to lower their fitness by behaving “for the good of the species” would quickly disappear from the population. Thus, gene-based theories of [p. 88 ↓] evolution were critical of group-selection ideas (Maynard Smith, 1964; Williams, 1966) and by the early 1970s, Wynne-Edwards and group selection were basically rejected by evolutionary biologists.

More recent developments in evolutionary biology now suggest that the original criticisms of group selection ideas were overstated. With the publication of Leo Buss's book, *The Evolution of Individuality* (1987), scientific consensus began shifting from gene-based selection models of evolution to multilevel evolutionary theories (Maynard Smith and Szathmáry, 1995; Sober and Wilson, 1998). L. Buss (1987) observed that biologists took the notion of the multicellular individual for granted. He argued that multicellularity itself evolved through the consolidation of initially self-replicating units. Evolutionary transitions creating new levels of selection involve both synergies and conflicts between lower and higher levels of organization. Multilevel evolutionary theory provided the needed conceptual frameworks for a new interpretation for the role of group selection in human evolution (Brewer and Caporael, 2006).

Multilevel or hierarchical models of evolution recognize that the concept of “fit” must be conceptualized in terms of embedded structures. Genes, as one level of organization, are adapted to fit the environment of their cellular machinery; cells fit the environment of the individual organism; and individual organisms are adapted to fit the next higher level of organization within which they function. Different levels of social organization and selection provide opportunities for both synergisms and conflicts between levels. In a hierarchical system, adaptive success at one level may need to be curtailed for the sake of success at a higher level in the system. Structural requirements at the higher level of organization constrain competition at lower levels.

This view of adaptation and natural selection provides a new perspective on the concept of group selection as a factor in human evolution (Brewer and Caprael, 2006; Caprael and Brewer, 1995). With coordinated group living as the primary survival strategy of the species, the social group, in effect, provided a buffer between the individual organism and the exigencies of the physical environment. As a consequence, then, the physical environment exercises only indirect selective force on human adaptation, while the requirements of social living constitute the immediate selective environment. The dynamics of multilevel selection resembles what Campbell (1974, 1990) called "downward causation" across system levels. Downward causation operates whenever structural requirements at higher levels of organization determine some aspects of structure and function at lower levels (a kind of reverse reductionism).

Both biological and behavioral scientists today accept the basic premise that human beings are adapted for group living. Even a cursory review of the physical endowments of our species – weak, hairless, extended infancy – makes it clear that we are not suited for survival as lone individuals, or even as small family units. Many of the evolved characteristics that have permitted humans to adapt to a wide range of physical environments, such as omnivorousness and tool making, create dependence on collective knowledge and cooperative information sharing. As a consequence, human beings are characterized by *obligatory interdependence* (Caprael and Brewer, 1995), and our evolutionary history is a story of co-evolution of genetic endowment, social structure, and culture.

If individual humans cannot survive outside of groups, then the structural requirements for sustaining groups create systematic constraints on individual biological and psychological adaptations. Cooperative groups must meet certain structural requirements in order to exist, just as organisms must have certain structural properties in order to be viable. For community-sized groups these organizational imperatives include mobilization and coordination of individual effort, communication, internal differentiation, optimal group size, and boundary definition. [p. 89 ↓] The benefits to individuals of cooperative arrangements cannot be achieved unless prior conditions have been satisfied that make the behavior of other individuals predictable and coordinated. Group survival depends on successful solution to these problems of internal organization and coordination. In other words, the viability of the group becomes a factor in the survival of individuals and their genetic reproduction. The

implication of this multilevel perspective on human evolution is that humans will be exquisitely sensitive to the viability of the groups they depend on (or commit themselves to), and that human motivation will be tuned to the requirements of the collective.

Optimal Distinctiveness Theory: Connecting the Threads

Optimal distinctiveness theory grew out of these three influences on my thinking about human sociality and group behavior. Social identity and ethnocentric ingroup bias suggested the importance of ingroup differentiation and intergroup distinctiveness in eliciting collective identity and concern for the welfare of others. The role of social identity in resolving social dilemmas defined conditions under which group welfare overrides individual self-interest and reinforced my conviction that a need for group identification is ‘built in’ to the human motivational system. Finally, the conceptual work on multilevel selection and group living provided an evolutionary framework for understanding that human nature is dualistic, and social motives reflect the tension between the requirements of individual and group survival.

Importantly, the development of optimal distinctiveness theory was in part the product of an exercise in thinking about downward causation from the group to the individual level of analysis. The advantage to early humans of extending social interdependence and cooperation to an ever wider circle of conspecifics comes from the ability to exploit resources across an expanded territory and buffer the effects of temporary depletions or scarcities in any one local environment. But expansion comes at the cost of increased demands on obligatory sharing and regulation of reciprocal cooperation. Both the carrying capacity of physical resources and the capacity for distribution of resources, aid, and information inevitably constrain the potential size of cooperating social networks. Thus, effective social groups cannot be either too small or too large. To function, social collectives must be restricted to some optimal size – sufficiently large and inclusive to realize the advantages of extended cooperation, but sufficiently exclusive to avoid the disadvantages of spreading social interdependence too thin.

Based on this analysis of one structural requirement for group survival, I hypothesized that the conflicting benefits and costs associated with expanding group size would have shaped social motivational systems at the individual level. If humans are adapted to live in groups and depend on group effectiveness for survival, then our motivational systems should be tuned to the requirements of group effectiveness. We should be uncomfortable depending on groups that are too small to provide the benefits of shared resources but also uncomfortable if group resources are distributed too widely. A unidirectional drive for inclusion would not have been adaptive without a counteracting drive for differentiation and exclusion. Opposing motives hold each other in check, with the result that human beings are not comfortable either in isolation or in huge collectives. These social motives at the individual level create a propensity for adhering to social groups that are both bounded and distinctive. As a consequence, groups that are optimal in size are those that will elicit the greatest levels of member loyalty, conformity, and cooperation, and the fit between individual psychology and group structure is better achieved.

In addition to representing the dual nature of human sociality, optimal distinctiveness [p. 90 ↓] theory was developed to fill a gap in extant theories of social identity. The original statements of social identity theory (Tajfel, 1981) and the subsequent development of self-categorization theory (Turner et al., 1987) were based heavily on cognitive processes of categorization and perceptual accentuation. This depiction provided an explanation for why and how specific social categorizations and ingroup–outgroup distinctions become salient but it lacks a driver for the process of *identification* with ingroups, particularly for chronic, long-term identification. Although the theory postulated that social identity salience had motivational *consequences* in the form of a striving for positive distinctiveness of the ingroup (Tajfel and Turner, 1979), a motivational component was missing from the theory with respect to antecedents of social identity.

For many social psychologists, the idea that social identification – with all its significant emotional and behavioral concomitants – is based solely on “cold cognition” was intuitively incomplete. Because group identity sometimes entails self-sacrifice in the interests of group welfare and solidarity, understanding why and when individuals are willing to relegate their sense of self to significant group identities requires motivational as well as cognitive analysis. Motivational explanations were also needed to account for

why group membership does not always lead to identification and why individuals are more chronically identified with some ingroups rather than others.

Optimal Distinctiveness: The Basic Model and Some Clarifications

Basic Premises of the Optimal Distinctiveness Model

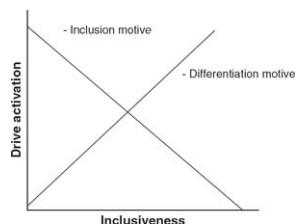
If social differentiation and intergroup boundaries are functional for social cooperation, and social cooperation is essential for human survival, then there should be psychological mechanisms at the individual level that motivate and sustain ingroup identification and differentiation. The optimal distinctiveness model (Brewer, 1991) posits that humans are characterized by two opposing needs that govern the relationship between the self-concept and membership in social groups. The first is a need for assimilation and inclusion, a desire for belonging that motivates immersion in social groups. The second is a need for differentiation from others that operates in opposition to the need for immersion. As group membership becomes more and more inclusive, the need for inclusion is satisfied but the need for differentiation is activated; conversely, as inclusiveness decreases, the differentiation need is reduced but the need for assimilation is activated. These competing drives hold each other in check, assuring that interests at one level are not consistently sacrificed to interests at the other. According to the model, the two opposing motives produce an emergent characteristic – the capacity for social identification with distinctive groups that satisfy both needs simultaneously.

The basic premise of the optimal distinctiveness model is that the two identity needs (inclusion/assimilation and differentiation/distinctiveness) are independent and work in opposition to motivate group identification. More specifically, it is proposed that social identities are selected and activated to the extent that they help to achieve a balance between needs for inclusion and for differentiation in a given social context.

Optimal identities are those that satisfy the need for inclusion *within* the ingroup and simultaneously serve the need for differentiation through distinctions *between* the ingroup and outgroups. In the original statement of the theory, I tried to capture the essential ideas in the form of a figure that depicted the opposing drives and the point of equilibrium ([Figure 30.1](#), adapted from Brewer, 1991).

In effect, optimal social identities involve *shared distinctiveness* (Stapel and Marx, [p. 91 ↓] 2007). Individuals will resist being identified with social categorizations that are either too inclusive or too differentiating but will define themselves in terms of social identities that are optimally distinctive. Equilibrium is maintained by correcting for deviations from optimality. A situation in which a person is overly individuated will excite the need for assimilation, motivating the person to adopt a more inclusive social identity. Conversely, situations that arouse feelings of deindividuation will activate the need for differentiation, resulting in a search for more exclusive or distinct identities.

Figure 30.1 The optimal distinctiveness model of group identification (from Brewer, 1991)



Evidence for competing social motives comes from empirical demonstrations of efforts to achieve or restore group identification when these needs are deprived. Results of experimental studies have shown that activation of the need for assimilation or the need for differentiation increases the importance of distinctive group memberships (Pickett et al., 2002), that threat to inclusion enhances self-stereotyping on group-characteristic traits (Brewer and Pickett, 1999; Pickett et al., 2002; Spears et al., 1997), and that threat to group distinctiveness motivates overexclusion (Brewer and Pickett, 2002) and intergroup differentiation (Roccas and Schwartz, 1993; Jetten et al., 1998; Hornsey and Hogg, 1999; Jetten et al., 2004). Further, assignment to distinctive minority group categories engages greater group identification and self-stereotyping than does

membership in large, inclusive majority groups (Brewer and Weber, 1994; Leonardelli and Brewer, 2001; Simon and Hamilton, 1994). Thus, there is converging evidence that group attachment is regulated by motives for inclusion and distinctiveness.

Some Qualifications and Clarifications

Although hypotheses derived from optimal distinctiveness theory have been tested by different researchers in many different contexts, some aspects of the theory are frequently misunderstood. Importantly, the model does *not* postulate that optimal distinctiveness is a property of some groups rather than others and that individuals directly seek identification with such optimal groups. Rather, optimality is an interactive product of current levels of activation of the opposing motives for inclusion and differentiation and group properties that determine its level of inclusiveness and distinctiveness. This leads to three important principles that are essential to understanding optimal distinctiveness.

First, *optimal distinctiveness is context specific*. Context affects both the activation [p. 92 ↓] of motives or needs and the relative distinctiveness of specific social categories. Consider, for example, my professional group memberships. In the context of an international psychology conference, categorization as a “psychologist” is far too inclusive, and a subcategory such as “social psychologist” is more likely to be optimally distinctive. On the other hand, in the context of my local community, my identity as a social psychologist is too highly differentiated and instead, categorizing my occupation as an “academic” is optimal. Identifying myself as an academic or a university professor places me in a social group with a significant number of other members of my community who share that occupational status and yet distinguishes us from neighbors who belong to other professions or occupational categories. “Shared distinctiveness” is contextually defined.

Second, *optimal distinctiveness is a dynamic equilibrium*. Even within a given context, optimality is not necessarily fixed because inclusion and differentiation motives are also subject to temporal influences and change over time. When one enters a new group, for example, the awareness of one's marginal status as a newcomer may enhance the need for inclusion relative to the need for differentiation, but as time

goes on and inclusion is more secure, differentiation needs become more salient and maintaining group distinctiveness assumes a higher priority. Groups also exhibit dynamic shifts across time in their relative focus on enhancing inclusiveness or reestablishing distinctiveness and exclusiveness.

Finally, *identity motives vary across situation, culture, and individuals*. Asking how “strong” an individual’s inclusion motive is like asking how strong is the individual’s hunger motivation. Like any need or drive, inclusion and differentiation motives vary as a function of current levels of satiation or deprivation. However, individuals may differ in how sensitive they are to changes in levels of inclusiveness. Just as some individuals start feeling ravenously hungry after an hour or two since they last ate while other individuals don’t even notice they haven’t eaten all day, so some people will react strongly to a slight loss of inclusiveness (or slight expansion of group boundaries), whereas others will be more tolerant of a range of ingroup inclusiveness. Thus, although the principles incorporated in the optimal distinctiveness model are presumed to be universal, the model can also accommodate individual, situational, and cultural differences in the relative activation of inclusion and differentiation needs and the nature of optimal identities.

Put more formally, the model as depicted in [Figure 30.1](#), has four important parameters: the height (intercept) of the need for differentiation, the height (intercept) of the need for inclusion, the negative slope of the need for inclusion, and the positive slope of the need for differentiation. Of these four, one is presumed to be fixed. The intercept (zero activation) of the need for differentiation is assumed to be at the point of complete individuation (the endpoint of the inclusiveness dimension). All of the other parameters are free to vary; any changes in the intercept or slope of the inclusion drive or the slope of the differentiation drive will alter the point of equilibrium that represents an optimal identity. Thus, the model depicted in [Figure 30.1](#) is just one member of a class of models containing all possible variations in these parameters, and differences across situations, cultures, and individuals can be represented in terms of variation in the slopes of the two drives (which can vary independently). (See Brewer and Roccas, 2001, for a discussion of how cultural differences can be reflected in model parameters and the point of equilibrium.) Again, the overall point is to emphasize that optimal distinctiveness is not a fixed property of groups or of individuals but a consequence of motivational dynamics at both levels.

[p. 93 ↓]

Implications of Optimal Distinctiveness Theory

Theoretical Implications

The theory of optimal distinctiveness (Brewer, 1991) is originally a theory of *collective* social identity. More specifically, it is a model of the opposing motives of inclusion and differentiation that regulate group identification, where a group is defined as a collective unit, or entity, that transcends individual level identities. The concept of opposing motives as a regulatory system, however, has potential broader application than just understanding group identification. I have suggested that analogous opposing needs for separateness and assimilation may also operate at the levels of individual and relational selves to determine optimal identities at those levels as well (Brewer and Gardner, 1996: 91; Brewer and Roccas, 2001, table 1). At the collective level, the conflict is between belonging and inclusion on the one hand, and separation and distinctiveness on the other. At the individual level, the needs are expressed in the opposition between the desire for similarity on the one hand and the need for uniqueness on the other (Snyder and Fromkin, 1980). At the interpersonal (relational) level, the tension is represented by conflicts between the need for autonomy and the need for interdependence and intimacy with specific others. At each level, the person must achieve some optimal balance between these conflicting motives for defining self in relation to others.

Social Implications: The Upside and the Downside of Optimal Distinctiveness

If social identity motives derive, ultimately, from needs for security and cooperative interdependence, this has important implications for the functions and limits of social

identification as a motivator of prosocial behavior. More specifically, the theory predicts that the dynamics of trust and cooperation will be shaped by the need for distinct boundaries between ingroup and outgroups and associated differences between intragroup and intergroup behaviors.

On the positive side, as I have noted, optimal group identities can be thought of as bounded communities of mutual trust and generalized reciprocal cooperation. Mere knowledge that another individual shares a salient group identity is sufficient to engage depersonalized trust, cooperative orientation, and willingness to sacrifice immediate self-interest for collective welfare. The dilemma in all this is that the conditions for ingroup cooperation and trust require group boundaries and clear differentiation between intragroup and intergroup social exchange. The social motives postulated by optimal distinctiveness theory at the individual level create a propensity for adhering to social groups that are both bounded and distinctive. Secure inclusion implies exclusion. The adaptive value of groups lies in interactional norms that facilitate reciprocal exchanges within the group, but are not extended to outsiders. A consequence of ingroup identification and intergroup boundaries is that individuals modify their social behavior depending on whether they are interacting with ingroup or outgroup members.

None of this implies that strong identification with ingroups necessitates conflict with outgroups. Contrary to the notion that ingroup positivity and outgroup derogation are reciprocally related, ingroup love does not imply outgroup hate (Brewer, 1999, 2001). What ingroup favoritism does imply is that positivity and trust extend only to the boundary of the ingroup and not across groups. Thus, intergroup relations are characterized by *lack* of trust, though not necessarily active *distrust*. For example, in our experiments on group-based trust, we find that knowing that a stranger belongs to one's own ingroup [p. 94 ↓] elevates trusting choices to near 90 percent (Yuki et al., 2005). With an outgroup stranger, on the other hand, trusting choices drop significantly – but only to around 50 percent, not to zero percent as would be expected if outgroups were assumed to be hostile and malevolent. Instead, exchanges with outgroup members appear to be characterized by uncertainty and lack of trust, rather than by automatic distrust or negativity.

Nonetheless, ingroup positivity and bounded trust are not completely benign. Just as there is a realistic basis for ethnocentric trust of ingroups, differences in norms

and sanctions applied to ingroup behavior compared to behavior in interactions with outgroup members provides a realistic basis for outgroup distrust and negative stereotypes. At the same time that groups promote trust and cooperation within, they caution wariness and constraint in intergroup interactions. Psychologically, expectations of cooperation and security promote positive attraction toward other ingroup members and motivate adherence to ingroup norms of appearance and behavior that assure that one will be recognized as a good or legitimate ingroup member. Symbols and behaviors that differentiate the ingroup from local outgroups become particularly important here, to reduce the risk that ingroup benefits will be inadvertently extended to outgroup members, and to ensure that ingroup members will recognize one's own entitlement to receive benefits. Assimilation within and differentiation between groups is thus mutually reinforcing, along with ethnocentric preference for ingroup interactions and institutions. Thus, even in the absence of overt conflict between groups, the differentiation between ingroup and outgroup behavior creates a kind of self-fulfilling prophecy in the realm of intergroup perceptions. As LeVine and Campbell (1972: 173) put it, "[I]f most or all groups are, in fact, ethnocentric, then it becomes an 'accurate' stereotype to accuse an outgroup of some aspect of ethnocentrism."

Combined with the accentuation principle that exaggerates perceived differences between social categories, this leads to a set of "universal stereotypes" to characterize ingroup–outgroup differences. Whereas "we" are trustworthy, peaceful, moral, loyal, and reliable; "they" are clannish, exclusive, and potentially untrustworthy. What is particularly interesting about this pattern of stereotypes is that the same behaviors that are interpreted as reasonable caution on the part of the ingroup in dealings with outgroup members become interpreted as "clannishness" and indicators of mistrust when exhibited by outgroupers toward the ingroup.

Although ingroup favoritism does not necessarily imply outgroup derogation, the motivational dynamics underlying strong ingroup attachment can lay the groundwork for intergroup hostility and conflict. Importantly, the critical function that ingroup distinctiveness holds for both survival of the group and individual psychological security explains why threats to identity can both engender and sustain strong intergroup conflict. Even in the absence of actual physical threat or material conflict of interest, the perception that the boundary between the ingroup and outgroup is being diluted or disrespected can create reactions equivalent to that of a territorial invasion. Bitter and

protracted conflict between different religious sects and ethnic subgroups testifies to the role of identity maintenance concerns even between subgroups within the context of a superordinate religion or nation. Especially in the modern world, competition over material resources such as land has as much to do with the identity meaning of those resources as it does actual group survival.

Optimal distinctiveness theory also has implications for when individuals will feel that their ingroup identity and the functions it serves are being threatened. If ingroups provide for both secure inclusion and intergroup differentiation, then anything that undermines either of these needs will activate attempts to restore optimality and enhance intergroup distinctions. The effects of threats to ingroup distinctiveness on hostility toward [p. 95 ↓] the threatening outgroup have been well documented (Jetten et al., 2004). But similar effects can be obtained when the individuals' sense of inclusion within the ingroup has been threatened (Pickett and Brewer, 2005). When a member of a group is led to believe that he or she is not a typical group member or is not fully accepted as part of the group, the person should experience distress to the extent that the person relies on that particular group for the satisfaction of belongingness, security, or assimilation needs. Peripheral group members not only need to be concerned with being similar to other ingroup members, but also concerned that they are not confused with the outgroup. This leads to the prediction that marginal ingroup members will be most concerned with maintaining intergroup distance and endorsing negativity toward outgroups.

In sum, then, understanding the origins and nature of ingroup favoritism, and differentiating ingroup attachment from outgroup hostility, may be critical for harnessing the best of human sociality while avoiding the consequences of intergroup hostility.

Conclusion

The dilemma that optimal distinctiveness theory poses for the modern world is this: How do we accommodate the need for distinctive ingroup identities that is rooted in our evolutionary past under conditions where interdependence transcends group boundaries at a global level? As a consequence of our evolutionary history, our sense of personal security and certainty are maximized in the context of shared ingroup

membership and clear ingroup–outgroup distinctions. The need for social identity and preservation of ingroup distinctiveness has long been recognized as a constraint on the “common ingroup identity” prescription for reducing intergroup discrimination and conflict. Attempts to merge groups or erase social category distinctions threaten optimal identities and limit our capacity for identification with larger, more inclusive categories.

It was this recognition that led Mummendey and Wenzel (1999) to argue that the question we should be asking is not how we can eliminate intergroup differences but rather under what conditions can intergroup differences be accepted, or even celebrated? The complexity of the modern world does provide us with multiple ways to meet identity needs, with multiple group identities that are optimal within different contexts. In a large and complex society, persons are differentiated or subdivided along many meaningful social dimensions, including gender and sexual orientation, life stage (e.g., student, worker, retiree), economic sector (e.g., technology, service, academics, professional), religion, ethnicity, political ideology, and recreational preferences. Each of these divisions provides a basis for shared identity and group membership that may become an important source of social identification. Further, most of these differentiations are cross-cutting in the sense that individuals may share a common ingroup membership on one dimension but belong to different categories on another dimension. Hence, having multiple group memberships has the potential to reduce the likelihood that one's social world can be reduced to a single ingroup–outgroup distinction. To the extent that we recognize the multiplicity and complexity of our own group identities, we may enhance the capacity for acceptance of intergroup differences and life in a pluralistic social system.

References

Brewer, M.B. Comment on Campbell's 'On the conflicts between biological and social evolution' . American Psychologist , 31, 372. (1976) <http://dx.doi.org/10.1037/0003-066X.31.5.372>

Brewer, M.B. In-group bias in the minimal intergroup situation: A cognitive motivational analysis . Psychological Bulletin , 86, 307–324. (1979) <http://dx.doi.org/10.1037/0033-2909.86.2.307>

Brewer, M.B. (1981) Ethnocentrism and its role in intergroup trust . In M. Brewer, ed. and B. Collins (eds), *Scientific Inquiry in the Social Sciences* , pp. 214–231. San Francisco: Jossey-Bass.

Brewer, M.B. (1986) The role of ethnocentrism in intergroup conflict . In W. Austin, ed. and S. Worchel, ed. and W. Austin (eds), *The Psychology of Intergroup Relations* , pp. 88–102. Chicago: Nelson-Hall.

Brewer, M.B. Ambivalent sociality: The human condition . *Behavioral and Brain Sciences* , 12, 699. (1989) <http://dx.doi.org/10.1017/S0140525X00025309>

Brewer, M.B. The social self: On being the same and different at the same time . *Personality and Social Psychology Bulletin* , 17, 475–482. (1991) <http://dx.doi.org/10.1177/0146167291175001>

Brewer, M.B. The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues* , 55, 429–444. (1999) <http://dx.doi.org/10.1111/0022-4537.00126>

Brewer, M.B. (2001) Ingroup identification and intergroup conflict: When does ingroup love become outgroup hate? In R. Ashmore, ed. , L. Jussim, ed. , and D. Wilder (eds), *Social Identity, Intergroup Conflict, and Conflict Reduction* . Oxford: Oxford University Press.

Brewer, M.B. and Campbell, D.T. (1976) Ethnocentrism and Intergroup Attitudes: East African Evidence . Beverly Hills: Sage.

Brewer, M.B. and Capra, L.R. Selfish genes versus selfish people: Sociobiology as origin myth . *Motivation and Emotion* , 14, 237–243. (1990) <http://dx.doi.org/10.1007/BF00996182>

Brewer, M.B. and Capra, L.R. (2006) An evolutionary perspective on social identity: Revisiting groups . In M. Schaller, ed. , J. Simpson, ed. , and D. Kenrick (eds), *Evolution and Social Psychology* , pp. 143–161. New York: Psychology Press.

Brewer, M.B. and Gardner, W. Who is this 'we'? Levels of collective identity and self representation . *Journal of Personality and Social Psychology* , 71, 83–93. (1996) <http://dx.doi.org/10.1037/0022-3514.71.1.83>

Brewer, M.B. and Kramer, R.M. Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing . *Journal of Personality and Social Psychology* , 50, 543–549. (1986) <http://dx.doi.org/10.1037/0022-3514.50.3.543>

Brewer, M.B. and Pickett, C.L. (1999) Distinctiveness motives as a source of the social self . In T. Tyler, ed. , R. Kramer, ed. , and O. John (eds), *The Psychology of the Social Self* , pp. 71–87. Mahwah, NJ: Lawrence Erlbaum and Associates.

Brewer, M.B. and Pickett, C.L. (2002) The social self and group identification: Inclusion and distinctiveness motives in interpersonal and collective identities . In J. Forgas, ed. and K. Williams (eds), *The Social Self: Cognitive, Interpersonal, and Intergroup Perspectives* , pp. 255–271. Philadelphia: Psychology Press.

Brewer, M.B. and Roccas, S. (2001) Individual values, social identity, and optimal distinctiveness . In C. Sedikides, ed. and M. Brewer (eds), *Individual Self, Relational Self, Collective Self* , pp. 219–237. Philadelphia: Psychology Press.

Brewer, M.B. and Schneider, S. (1990) Social identity and social dilemmas: A double-edged sword . In D. Abrams, ed. and M. Hogg (eds), *Social Identity Theory: Constructive and Critical Advances* , pp. 169–184. London: Harvester-Wheatsheaf.

Brewer, M.B. and Weber, J.G. Self-evaluation effects of interpersonal versus intergroup social comparison . *Journal of Personality and Social Psychology* , 66, 268–275. (1994) <http://dx.doi.org/10.1037/0022-3514.66.2.268>

Buss, L.W. (1987) *The Evolution of Individuality* . Princeton: Princeton University Press.

Campbell, D.T. (1969) Ethnocentrism of disciplines and the fish-scale model of omniscience . In M. Sherif, ed. and C. Sherif (eds), *Interdisciplinary Relationships in the Social Sciences* . Hawthorne, NY: Aldine.

Campbell, D.T. (1974) 'Downward causation' in hierarchically organised biological systems . In F. Ayala, ed. and T. Dobzhansky (eds), *Studies in the Philosophy of Biology* , pp. 179–186. London: Macmillan.

Campbell, D.T. On the conflicts between biological and social evolution and between psychology and moral traditions . *American Psychologist* , 30, 1103–1126. (1975) <http://dx.doi.org/10.1037/0003-066X.30.12.1103>

Campbell, D.T. (1990) Levels of organization, downward causation, and the selection-theory approach to evolutionary epistemology . In G. Greenberg, ed. and E. Tobach (eds), *Theories of the Evolution of Knowing* , pp. 1–17. Hillsdale, NJ: Erlbaum.

Capra, L.R. and Brewer, M.B. Hierarchical evolutionary theory: There is an alternative, and it's not creationism . *Psychological Inquiry* , 6, 31–34. (1995) http://dx.doi.org/10.1207/s15327965pli0601_2

Capra, L.R., Dawes, R.M., Orbell, J.M. and van de Kragt, A. Selfishness examined: Cooperation in the absence of egoistic incentives . *Behavioral and Brain Sciences* , 12, 683–739. (1989) <http://dx.doi.org/10.1017/S0140525X00025292>

De Cremer, D. and van Dijk, E. Reactions to group success and failure as a function of identification level: A test of the goal-transformation hypothesis in social dilemmas . *Journal of Experimental Social Psychology* , 38, 435–442. (2002) <http://dx.doi.org/10.1016/S0022-1031%2802%2900009-4>

De Cremer, D. and Van Vugt, M. Social identification effects in social dilemmas: A transformation of motives . *European Journal of Social Psychology* , 29, 871–893. (1999) <http://dx.doi.org/10.1002/%28SICI%291099-0992%28199911%2929:7%3C871::AID-EJSP962%3E3.0.CO;2-I>

Foddy, M., Platow, M. and Yamagishi, T. (2008) Group-based trust in strangers: The role of stereotypes and group heuristics . Unpublished manuscript.

Hardin, G. The tragedy of the commons . *Science* , 162, 1243–1248. (1968) <http://dx.doi.org/10.1126/science.162.3859.1243>

Hornsey, M.J. and Hogg, M.A. Subgroup differentiation as a response to an overly-inclusive group: A test of optimal distinctiveness theory . European Journal of Social Psychology , 29, 543–550. (1999) <http://dx.doi.org/10.1002/%28SICI%291099-0992%28199906%2929:4%3C543::AID-EJSP945%3E3.0.CO;2-A>

Jetten, J., Spears, R. and Manstead, A.S.R. Intergroup similarity and group variability: The effects of group distinctiveness on the expression of in-group bias . Journal of Personality and Social Psychology , 74, 1481–1492. (1998) <http://dx.doi.org/10.1037/0022-3514.74.6.1481>

Jetten, J., Spears, R. and Postmes, T. Intergroup distinctiveness and differentiation: A meta-analytic integration . Journal of Personality and Social Psychology , 86, 862–879. (2004) <http://dx.doi.org/10.1037/0022-3514.86.6.862>

Kramer, R.M. and Brewer, M.B. Effects of group identity on resource utilization in a simulated commons dilemma . Journal of Personality and Social Psychology , 46, 1044–1057. (1984) <http://dx.doi.org/10.1037/0022-3514.46.5.1044>

Kramer, R.M. and Brewer, M.B. (1986) Social group identity and the emergence of cooperation in resource conservation dilemmas . In H. Wilke, ed. , D. Messick, ed. , and C. Rutte (eds), *Psychology of Decisions and Conflict*, Vol. 3: Experimental Social Dilemmas , pp. 205–230. Frankfurt: Verlag Peter Lang.

Kramer, R.M. and Wei, J. (1999) Social uncertainty and the problem of trust in social groups: The social self in doubt . In T.R. Tyler, ed. , R.M. Kramer, ed. , and O.P. John (eds), *The Psychology of the Social Self* , pp. 145–168. Mahwah, NJ: Lawrence Erlbaum Associates.

Leonardelli, G. and Brewer, M.B. Minority and majority discrimination: When and why . Journal of Experimental Social Psychology , 37, 468–485. (2001) <http://dx.doi.org/10.1006/jesp.2001.1475>

LeVine, R.A. and Campbell, D.T. (1972) *Ethnocentrism: Theories of Conflict, Ethnic Attitudes and Group Behavior* . New York: Wiley.

Maynard Smith, J. Group selection and kin selection . *Nature* , 201, 1145–1147. (1964) <http://dx.doi.org/10.1038/2011145a0>

Maynard Smith, J. and Szathmáry, E. (1995) *The Major Transitions in Evolution* . Oxford: W.H. Freeman.

Mummendey, A. and Wenzel, M. Social discrimination and tolerance in intergroup relations: Reactions to intergroup difference . *Personality and Social Psychology Review* , 3, 158–174. (1999) http://dx.doi.org/10.1207/s15327957pspr0302_4

Parker, R., Lui, L., Messick, D.M., Messick, C., Brewer, M.B., Kramer, R., Samuelson, C. and Wilke, H. A computer laboratory for studying resource dilemmas . *Behavioral Science* , 28, 298–304. (1983) <http://dx.doi.org/10.1002/bs.3830280406>

Pickett, C.L., Bonner, B.L. and Coleman, J.M. Motivated self-stereotyping: Heightened assimilation and differentiation needs result in increased levels of positive and negative self-stereotyping . *Journal of Personality and Social Psychology* , 82, 543–562. (2002) <http://dx.doi.org/10.1037/0022-3514.82.4.543>

Pickett, C.L. and Brewer, M.B. (2005) The role of exclusion in maintaining in-group inclusion . In D. Abrams, ed. , M. Hogg, ed. , and J. Marques (eds), *Social Psychology of Inclusion and Exclusion* , pp. 89–112. New York: Psychology Press.

Pickett, C.L., Silver, M.D. and Brewer, M.B. The impact of assimilation and differentiation needs on perceived group importance and judgments of group size . *Personality and Social Psychology Bulletin* , 28, 546–558. (2002). <http://dx.doi.org/10.1177/0146167202287011>

Roccas, S. and Schwartz, S. Effects of intergroup similarity on intergroup relations . *European Journal of Social Psychology* , 23, 581–595. (1993) <http://dx.doi.org/10.1002/ejsp.2420230604>

Simon, B. and Hamilton, D.L. Social identity and self-stereotyping: The effects of relative group size and group status . *Journal of Personality and Social Psychology* , 66, 699–711. (1994) <http://dx.doi.org/10.1037/0022-3514.66.4.699>

Snyder, C.R. and Fromkin, H.L. (1980) Uniqueness: The Human Pursuit of Difference . New York: Plenum Press. <http://dx.doi.org/10.1007/978-1-4684-3659-4>

Sober, E. and Wilson, D.S. (1998) Unto Others: The Evolution and Psychology of Unselfish Behavior . Cambridge, MA: Harvard University Press.

Spears, R., Doosje, B. and Ellemers, N. Self-stereotyping in the face of threats to group status and distinctiveness: The role of group identification . Personality and Social Psychology Bulletin , 23, 538–553. (1997) <http://dx.doi.org/10.1177/0146167297235009>

Stapel, D.A. and Marx, D.M. Distinctiveness is key: How different types of self-other similarity moderate social comparison effects . Personality and Social Psychology Bulletin , 33, 437–448. (2007) <http://dx.doi.org/10.1177/0146167206296105>

Sumner, W.G. (1906) Folkways . New York: Ginn.

Tajfel, H. Experiments in intergroup discrimination . Scientific American , 223, 96–102. (1970) <http://dx.doi.org/10.1038/scientificamerican1170-96>

Tajfel, H. (1981) Human Groups and Social Categories . Cambridge: Cambridge University Press.

Tajfel, H., Billig, M., Bundy, R. and Flament, C. Social categorization and intergroup behaviour . European Journal of Social Psychology , 1, 149–178. (1971) <http://dx.doi.org/10.1002/ejsp.2420010202>

Tajfel, H., and TurnerJ.C. (1979) An integrative theory of intergroup conflict . In W. Austin, ed. and S. Worchel (eds), Social Psychology of Intergroup Relations , pp. 33–47. Chicago, Nelson.

Tanis, M. and Postmes, T. A social identity approach to trust: Interpersonal perception, group membership and trusting behaviour . European Journal of Social Psychology , 35, 413–424. (2005) <http://dx.doi.org/10.1002/ejsp.256>

Turner, J.C., Hogg, M., Oakes, P., Reicher, S. and Wetherell, M. (1987) Rediscovering the Social Group: A Self-categorization Theory . Oxford: Blackwell.

Williams, G.C. (1966) *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought* . Princeton, NJ: Princeton University Press.

Wilson, E.O. (1975) *Sociobiology* . Cambridge, MA: Harvard University Press.

Wynne-Edwards, V.C. (1962) *Animal Dispersion in Relation to Social Behaviour* . London: Oliver and Boyd.

Yuki, M., Maddux, W.W., Brewer, M.B. and Takemura, K. Cross-cultural differences in relationship- and group-based trust . *Personality and Social Psychology Bulletin* , 31, 48–62. (2005) <http://dx.doi.org/10.1177/0146167204271305>

<http://dx.doi.org/10.4135/9781446249222.n30>

Handbook of Theories of Social Psychology

A Cognitive-Neoassociation Theory of Aggression

Contributors: Leonard Berkowitz

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "A Cognitive-Neoassociation Theory of Aggression"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n31>

Print pages: 99-119

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n31>

[p. 99 ↓]

Chapter 31: A Cognitive-Neoassociation Theory of Aggression

Leonard Berkowitz

Abstract

The author's cognitive-neoassociation (CNA) analysis of impulsive aggressive reactions can be traced back to the 1939 Dollard et al. frustration – aggression hypothesis, and to its later extensions by Neal Miller, such as his conflict model of hostility displacement. The author's research and writings, starting in the late 1950s, have been generally sympathetic to this perspective. Much of this research initially had to do with the aggression-enhancing influence of situational stimuli, such as weapons and movie violence, but increasingly, starting in the mid 1980s, particular attention was given to the role of negative affect. The author has modified the original frustration – aggression hypothesis by proposing that obstacles to expected goal attainment produce aggressive inclinations only to the extent that these events are experienced as decidedly unpleasant. In spelling out the CNA model, the present chapter maintains that aggression-related stimuli and aversive occurrences tend to activate aggressive reactions automatically and that cognitive processing can then intervene to enhance or weaken the aggressive inclinations. The chapter then concludes by raising a number of important questions still to be resolved.

Theoretical Foundation

The theoretical perspective I have employed throughout my research on aggression has a clear starting point: the publication of the monograph *Frustration and Aggression* in 1939 by John Dollard, Leonard Doob, Neal Miller, O. Hobart Mowrer, and Robert

Sears, then all at Yale University. Here, in this relatively small book, the authors argued that “aggression is always a consequence of frustration” (1939: 1), and extended this central proposition very broadly to such matters as socialization, adolescent behavior, criminality, and even fascism and communism. This publication drew so much attention throughout psychology at that time that two-thirds of a 1941 issue of the *Psychological Review* was devoted to discussions of this analysis, and these discussion papers were then reprinted in full in the then singularly important *Readings in Social Psychology* published in 1947.¹

I was so attracted to this conception, when I joined the Psychology Department at the [p. 100 ↓] University of Wisconsin in 1955 and was asked to teach a senior-level course on special topics in psychology, that I devoted most of the semester to research on aggression and centered this coverage on the ideas advanced by the Yale group. My lecture notes then served as the foundation for a successful research grant application to the National Institute of Mental Health and also for an article in the *Psychological Bulletin* (Berkowitz, 1958). A few years later I (Berkowitz, 1963) published a comprehensive survey of the psychological research on aggression, again sympathetic to the perspective employed by Dollard, Miller, and their colleagues and its later modifications and extensions (Miller, 1941, 1948; also see Miller, 1959). But even with all this attention to the frustration – aggression hypothesis in my teaching and literature reviews, my own early research in the 1960s and 1970s did not directly examine the effects of frustrations on subsequent aggression, although I did return to this topic in an edited volume (Berkowitz, 1969) and in a 1989 survey of the pertinent literature (Berkowitz, 1989). My student, Russell Geen (1968), did study the influence of frustrations in his doctoral research. Other problems, such as the displacement of hostility, were then more interesting to me, and I'll say more about this later.

My interest in the Yale group's theoretical perspective does not mean I agreed with this analysis in all respects. Where Dollard and his colleagues (1939) had originally held that every aggressive action “presupposes the existence of frustration” (1939: 1), it seemed obvious to me (e.g., see Berkowitz, 1963, 1989), that aggression can occur at times when there is no barrier to goal attainment, and can have objectives in addition to the target's injury. Many attacks are probably primarily instrumental to the attainment of

some nonaggressive goal, and the Dollard et al. (1939) analysis holds only for a limited range of situations.

With this qualification, the Yale writers' theorizing appealed to me for a number of reasons, most obviously for its sweep and its testability, and not for its central postulate – but also because their thinking was more complex than is commonly realized. These writers did not hold that aggression was the only or even the main response to a thwarting, as Miller (1941) carefully noted. They suggested that an impediment to goal attainment produced instigations to a variety of different actions, of which the instigation to aggression was only one. However, Miller (1941) also said, if the nonaggressive behaviors did not successfully remove the frustration, "the greater is the probability that the instigation to aggression eventually will become dominant so that some responses of aggression actually will occur" (1941: 339).

Theoretical Conceptions Derived from the Yale Group's Analyses

Automatic and/or Cognitively Controlled Aggressive Responses

Yet another reason I liked the Yale group's approach was their implicit conception of the frustration reactions as being automatically evoked. Because of the writers' general adherence to Hullian behavior theory (see Miller, 1959), I assumed they believed the thwarting-produced reactions would occur automatically – with little thought and attention and not guided by any intentions other than the urge to hurt the target. I emphasized this automaticity in my own thinking about many aggressive actions, especially, but not only, those carried out in a fit of rage (e.g., see Berkowitz, 2008). However appealing the Yale group's theorizing was to me, nonetheless, it didn't match the view of aggression widely shared throughout the social sciences. Most analyses of aggression in these disciplines basically think attacks result when the perpetrators decide, not necessarily consciously, that their purposes can be well [p. 101 ↓] served

by hurting the target. Script theory is a contemporary version of this perspective; it essentially contends that the aggressors, following the cognitive scripts in their minds suggesting what is likely to happen in the presenting situation, choose to assault their target (e.g., Bushman and Anderson, 2001; Huesmann, 1988).

Is the Frustration–Aggression Relationship the Result of Learning?

It also seemed to me that the Dollard et al. (1939) formulation suggested that the frustration-produced aggressive reactions were, at least to some extent, the result of an innate process, although Miller had stated that he and his colleagues had made no assumptions “as to whether the frustration–aggression relationship is of innate or of learned origin” (1941: 340). Still, many critics caught the implication of a “built-in” basis to the presumed connection between thwartings and aggressive reactions. Maintaining that frustration reactions were learned, they held that frustrations did not necessarily produce an aggressive urge (e.g., Bandura and Walters, 1963). Experiments with infants, however, indicate that quite a few very young children display facial expressions indicative of anger when they are frustrated by the unexpected removal of either a pleasant picture (Lewis, 1993) or a desired toy (Stifter and Grant, 1993), so that it is indeed possible that angry reactions to thwartings are not necessarily always learned.

Hostility Displacement

After discussing the relatively indirect forms of aggression that theoretically would occur when direct attacks on the frustrater are inhibited, Dollard and his colleagues (1939) pointed out that the restrained aggressive urge might also be expressed in attacks on persons other than the aggression instigator. Explicitly adopting Freudian terminology, they referred to this phenomenon as aggression displacement (1939: 41). Social psychology was once quite interested in displaced aggression, as Marcus-Newhall and her colleagues (2000) observed after content analyzing a great many social psychology texts. In recent years, however, according to Marcus-Newhall et al. (2000), displaced aggression has been given little attention by psychological investigators, although, as

this survey concluded, it is a “robust” phenomenon. It’s worth looking at aggression displacement again since it is very much in accord with my theoretical perspective.

Aggression Displacement in Hostility toward Minorities

In their discussion of the displacement of aggression (e.g., 1939: 41–44), the Yale writers noted that this displacement could also be manifested in hostile attitudes toward minority groups such as “Negroes.” The Hovland and Sears (see Dollard et al., 1939: 31) investigation is undoubtedly the best known of their studies on this topic. Because cotton was the main cash crop in the South at that time, these researchers assumed that Southern farmers suffered economic-related frustrations when cotton prices were low. In keeping with their expectation, cotton value in Southern states between 1882 and 1930 was significantly negatively correlated with the number of lynchings of blacks in these states in those years. The farmers’ economic hardships apparently had produced aggressive urges which were then displaced onto blacks.

This study was widely discussed in the social sciences in the succeeding decades, sometimes drawing criticism, but also support from more sophisticated statistical analyses (see Green et al., 1998). In what is the most thorough of these follow-up investigations, Green et al. (1998) showed that the cotton price/lynching relationship held only for the time up to the Great Depression, but not afterward. And moreover, on extending their investigation to the [p. 102 ↓] effects of economic difficulties on nonlethal hate crimes, Green et al. (1998) found that unemployment rates in New York City in the decade before their study had no relationship to the number of reported hate crimes against homosexuals, Jews, blacks, and Asians during that period. All in all, it could be that when people experience economic frustrations and/or other social stresses, they will openly direct their resulting aggressive urges onto particular minority groups only if they think others important to them, their ingroups, will not disapprove of these assaults. Inhibitions restraining direct attacks against minorities also could have increased substantially in recent decades, dampening overt displays of bigotry. It is even conceivable that minority groups as such no longer possess the decidedly negative stimulus qualities that draw assaults from those disposed to be violent.

Stimulus Qualities Drawing Displaced Aggression

Neal Miller's (1948) seminal demonstration of how hostility displacement can be understood in stimulus-response generalization terms can help explain why some available targets are attacked and others are not victimized. Often called a conflict model because it deals with situations in which an aggressor both desires and fears to attack someone openly, Miller's (1948) analysis proposes that how the conflict is resolved depends upon three factors: the strength of this aggressive instigation (often generally termed the approach tendency), the strength of whatever instigations there are at the time to inhibit the open display of direct aggression (the avoidance tendency), and the degree of association between each possible target and the original provocateur. Miller (1948, 1959) also proposed that the avoidance tendency (the inhibitory generalization gradient) frequently mounts more rapidly than does the approach tendency (the instigation to attack generalization gradient) the closer the association between each target and the provocateur. And so, the model says, thwarted people wanting to assault their frustraters, but who are afraid to do so, will refrain from attacking the provocateurs or perhaps even others who are closely associated with the frustraters, but will instead direct their aggression toward those moderately linked to the source of their disturbance. The possible targets having little or no connection with the provocateur will receive little, if any, aggression.

My reanalysis of Fitz's (1976) experiment documents the applicability of Miller's (1948) model. In this study, angered men who believed they could safely "get even" with the person who had insulted them showed the pattern Miller had predicted: they attacked the provocateur most severely, and another individual associated with him next most intensely, whereas a nonassociated stranger received the lowest level of punishment. And also in accord with the model, the men led to be afraid of the insulter's possible retaliation also exhibited the displaced aggression pattern; here the insulter received relatively little punishment, but the provocateur's friend was punished much more severely (and the stranger was given little punishment).

Various Associations Linking the Available Target to the Anger Instigator

Aggression-eliciting associations can be established in a variety of ways. Hewitt (1974) showed that aggression can be generalized on the basis of the available target's similarity in age to the angering source. Another experiment, conducted by Berkowitz and Knurek (1969), indicated that having the same negative label can also connect the anger instigator to someone else. In this study the participants were provoked by a person having a name they earlier had been conditioned to dislike. When they later interacted with a peer having either the negatively conditioned name or a neutral name, they were harsher in their evaluation of the individual with the unpleasant name than the [p. 103 ↓] peer with the neutral name. The hostility generated by the negatively named provocateur had evidently generalized to the person bearing the same unpleasant label.

Dislike for the Minority Group – A Factor in Scapegoating?

I have long maintained that we have to consider a minority group's stimulus qualities if we are to explain why some minorities are especially apt to be the targets of displaced aggression. Although theorists have accounted for this selectivity in various ways (see Brewer and Brown, 1998), in my view most of the points they make can be subsumed under one significant general principle: They all provide reasons why the particular group is greatly disliked; that is, has acquired a strongly negative cue value. As a consequence, I suggest, the hostility aroused by other sources can be readily generalized to that particular collection of people.

A number of experiments have reported findings consistent with this proposition. In one of the earliest of these studies (Berkowitz and Holmes, 1960), the female participants were first induced to either dislike or have a neutral attitude toward a peer and then were either insulted or treated in a neutral manner by the experimenter. When all of the women were then given an opportunity to deliver electric shocks to their peer,

supposedly as an evaluation of her work on a task, those who had been provoked by the experimenter administered the severest punishment to the person they had earlier learned to dislike. This latter individual's negative cue value apparently had enhanced her ability to draw the hostility engendered by the provocateur.

Another Wisconsin study suggests that people highly disposed to be prejudiced are especially apt to exhibit this hostility generalization to disliked persons. This experiment (Berkowitz, 1959) took advantage of the freedom many Midwestern college students felt at that time to express prejudiced opinions openly. After the female participants had been deliberately derogated by the experimenter, those who had highly anti-Semitic attitudes tended to be the most hostile toward a neutral woman nearby.²

The Cognitive-Neoassociation Perspective

In the decades following the studies just cited, my emphasis on the role of situational stimuli automatically eliciting aggressive reactions continued, but I also developed some new conceptions not anticipated by the traditional S-R perspective. One of these is an important revision of the frustration – aggression hypothesis, and another has to do with the interplay of automatic and controlled cognitive processing in the display of aggressive actions.

Why Frustrations Produce Aggressive Reactions: The Role of Negative Affect

My 1989 review of the frustration–aggression research (Berkowitz, 1989) offered a possible explanation for why people do not always want to attack someone after they've been thwarted in their attempt to reach a desired goal: they are not sufficiently bothered. Barriers to goal attainment, I proposed (also see Berkowitz, 1983, 1993), produce an instigation to aggression only to the extent that they are decidedly unpleasant.

From this perspective, unexpected or unjustified interferences are more apt to provoke an aggressive reaction than anticipated or legitimate barriers to goal attainment

because the former are usually much more unpleasant. And similarly, another person's deliberate attempt to block our goal attainment is more angering than an inadvertent impediment because the former frustration is more disturbing. Furthermore, the factors identified by Dollard and his colleagues as determining the strength of the [p. 104 ↓] frustration-produced instigation to aggression, such as the intensity of the drive that cannot be satisfied or the extent to which goal attainment is blocked (Dollard et al., 1939: 28), have this effect because they govern the magnitude of displeasure that is experienced. My formulation also holds that the aggression-instigating effects of frustrations and insults cannot be compared in the abstract, as some psychologists have done. All frustrations are not equally bothersome, and all insults do not generate the same displeasure. In sum, it is not the exact nature of the aversive incident that is important but how intense is the resulting negative affect.

The Cognitive-Neoassociationistic Perspective

My theoretical analysis, which I call a cognitive-neoassociationistic (CNA) model, obviously was influenced to a great extent by the learning theory/associationistic theorizing prominent in psychology before the "cognitive revolution" of the 1960s. But it also was shaped to a large degree by Bower's cognitive-neoassociationism, especially his studies of the effects of mood on memory (e.g., Bower, 1981).

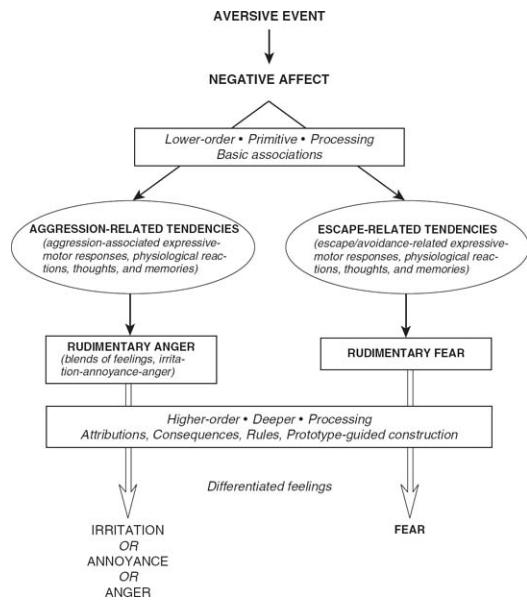
My version of this line of thought holds that both cognitive and automatic, associative processes can function to evoke largely involuntary aggressive reactions. The cognitions in this case can define an event as decidedly unpleasant (although of course, some occurrences are aversive in themselves) and they can also operate to impart an aggressive meaning to situational details. But CNA is primarily concerned with the consequences of aversive stimuli, however this aversiveness originates. Once the strong displeasure is experienced, the model says, there is a sequence of responses. The reactions initially are largely governed by associative processes, with cognitions presumably becoming more important in the later stages. Basically, as can be seen in [Figure 31.1](#), decidedly unpleasant conditions initially tend to activate, automatically and with relatively little thought, at least two sets of "primitive" inclinations: one to escape

from or avoid the aversive stimulation, *and also*, another to attack and even destroy the source of this stimulation. In other words, the aversive state of affairs presumably gives rise to *both* flight *and* fight tendencies. Neither of these inclinations is always dominant. Genetic factors, prior learning, and situational influences all enter to determine the relative strengths of these various reactions.

Also important, both the flight and fight tendencies should be regarded as syndromes, networks of associatively linked physiological, motoric, and cognitive components. The activated flight-associated syndrome is consciously experienced as *fear*, whereas the activated fight-linked syndrome is felt as *annoyance* or *irritation* (at relatively weak levels) or *anger* (at more intense levels). Because the syndromes are associatively linked networks, the activation of any particular syndrome component will also tend to bring the other parts of the network into operation. Thus, more than the formulations resting exclusively on cognitive concepts, this theory can accommodate the research showing that the display of the facial expressions and bodily postures characteristic of a given affective state, such as anger, can generate the feelings typical of this state (see Duclos et al., 1989).

CNA differs from the Anderson's general aggression model (e.g., see Anderson et al., 1995) primarily through its emphasis on these first, relatively automatic and noncognitively mediated reactions to the experienced negative affect. But recognizing the important role cognitions can play, CNA also proposes, as Figure 31.1 shows, that the first fairly primitive reactions can be modified and even substantially altered after the first response tendencies arise. In this second phase any active information processing can bring appraisals, attributions, and the like, into operation, thereby modifying or extinguishing the initial reactions. Thus, people can become angry and assault someone [p. 105 ↓] impulsively at times without the intervention of the complicated thought processes postulated by appraisal/attribution theorizing. As a matter of fact, I suggest that people exposed to decidedly unpleasant stressful conditions sometimes blame a salient available target for their troubles because of the hostile thoughts and angry feelings that had been generated within them; their attributions might then be the result rather than the cause of their affective reactions (see Quigley and Tedeschi, 1996, and Keltner et al., 1993, for supporting evidence).

Figure 31.1 The cognitive-neoassociationistic model



Evidence consistent with CNA can be found in the investigations demonstrating that exposure to decidedly unpleasant physical stimuli increases the chances of aggressive behavior (see, among other sources, Berkowitz, 1983, 1993). My thinking here isn't entirely new (e.g., see Baron et al., 1974), but along with Anderson (e.g., Anderson et al., 1995), I have made this a major proposition in my CNA analysis of affective aggression, and contend that negative affect will evoke a number of negative reactions including an instigation to aggression.

[p. 106 ↓]

Effects of Physically Unpleasant Conditions

Physical Pain

Physical pain obviously is almost always aversive and it often instigates aggressive inclinations. On studying patients suffering from frequent bouts of pain, Fernandez and Turk (1995: 165) maintained that anger is a "feature of chronic pain" and that

"anger stands out as one of the most salient emotional correlates of pain" (also see Greenwood et al., 2003). Of course, the anger could well contribute to the heightened pain (Fernandez and Turk, 1995; Greenwood et al., 2003), but the pain might also produce anger and aggressive inclinations. Fernandez and Wasan (in press) have recognized this possibility by noting that the suffering persons often develop hostile appraisals of those around them. Couldn't these appraisals be manifestations of the anger and aggressive urges they are experiencing?

Research conducted by Berkowitz et al. (1981) indicates that even relatively moderate pain levels can instigate aggressive impulses. The pain inflicted in these investigations was produced by having half of the female participants immerse a hand in very cold water for about six minutes as they evaluated a partner's solutions to several assigned problems. For the other women the water was at a more comfortable room temperature as they carried out their evaluations. In providing their assessments the participants could give their partner either rewards (nickels) or punishments (unpleasant noise blasts). Cutting across the water temperature variation, half of the women were told any punishment they delivered would actually help the problem solver by motivating her to think better, whereas the remaining women were informed the punishment would hurt their partner.

Those experiencing the moderate pain and who believed they could hurt their partner tended to deliver the most punishments relative to the rewards they provided. Even this mild pain level had evidently heightened their urge to hurt the available target.

Unpleasant Atmospheric Conditions

Decidedly unpleasant environmental conditions can also increase the proclivity to aggression even though they are not clearly painful. Here is a very brief and limited look at an extensive and controversial literature dealing with the effects of high ambient temperatures (see, for example, Anderson and Anderson, 1996, 1998; Cohn and Rotton, 1997, for more complete discussions).

Social scientists have long noted that, in a number of countries, violent crime rates tend to be higher in their warmer, southern latitudes than in the cooler, northern regions.

Anderson and Anderson (1996, 1998) added to these early observations by citing more recent studies showing essentially the same kind of south–north differences. Their own sophisticated analysis of area differences within the US demonstrated that cities having the hottest weather typically had the highest violence rates even when their social and economic characteristics were partialled out (1998: 264–265).

The relatively high homicide rate in the southern US is a good example of such a regional effect on violence. Over the generations, more murders and assaults have been committed in the southern states, controlling for population size, than in the northern parts of the country. Nisbett and Cohen (1996), among others, have attributed this difference primarily to a prevailing *culture of honor* in the South. White males growing up in this area presumably have learned that they must redress a perceived threat to their honor by attacking the offender in order to protect their image as tough and able to protect themselves and their possessions. Although Nisbett and Cohen (1996) provided evidence consistent with their thesis, Anderson and Anderson (1996, 1998) argued that regional temperature is a better predictor of the South's high violence rates than is its “culture of honor.” When they analyzed data from 260 standard metropolitan areas in the US for 1980, they concluded that the southerners could have developed and maintained their violence-encouraging attitudes and [p. 107 ↓] values largely because of the region's hot weather (1998: 270).

Other research shows that it is important to consider the likelihood and nature of interpersonal encounters when explaining the relationship between temperature and violent offenses. Data obtained by Cohn and Rotton (1997) indicated that the persons exposed to unpleasantly high temperatures might well have an instigation to aggression, but if they are engaged in distracting activities and/or if there is no suitable target nearby and/or if their inhibitions against aggression are fairly strong because of the presence of nearby aggression-restraining people, their urge will not become manifest in open behavior.

The findings from relevant laboratory experiments seem quite inconsistent (see Anderson and Anderson, 1998: 283). The Anderson group's studies repeatedly showed that an unpleasant temperature (cold or hot) produced stronger negative affect, angry feelings, and hostile attitudes and cognitions than did more comfortable ambient temperatures. But these reactions were not always accompanied by strong attacks on

an available target. In one of their investigations (see Anderson and Anderson, 1998), as an example, the participants in either an uncomfortably cold or uncomfortably hot room were angrier and had more hostile attitudes than their counterparts exposed to more moderate temperatures – but the former were more aggressive to a partner only the first time they could administer punishment and not on the later occasions. Presumably being aggressively inclined, it may be that they attacked their competitor more or less impulsively at first, but then may have decided it was best to restrain themselves.

Social Stresses

A number of social science theorists have proposed that harsh social situations, and especially frustrations, are major contributors to criminal activities. Sociologists and criminologists often speak of these conditions as *social strains* and refer to *strain theories* of crime causation, whereas psychologists typically employ the term *frustration* much more explicitly. But whatever words are used in these analyses, they highlight the central role of negative affect in antisocial behavior.

An experiment by Passman and Mulhern (1977) is relevant. The mothers in this study worked on an assigned task at the same time that they monitored their child's performance on a puzzle. Those women who were under a high degree of stress as they worked, because their task requirements had been deliberately made confusing, punished their youngsters' mistakes more severely than did their more comfortable counterparts. The stresses encountered in the everyday world can lead to more naturalistic aggression as well. Straus (1980) and his colleagues asked the men and women in their nationally representative US sample to indicate whether they had experienced each of 18 stressful life events – such as "troubles with other people at work," "the death of someone close," "a move to a different neighborhood or town," and "a family member with a health or behavior problem." Whether the respondents were male or female, the greater were the number of stressors they reported experiencing during the past year, the more likely they were also to say they had abused their children. Going further along these lines, other investigations have indicated that the psychological pain felt on being socially rejected can produce aggressive reactions (MacDonald and Leary, 2005) and that the negative affect some people experience at

the sight of homosexual activity can prompt aggressive inclinations toward gay men (Parrott et al., 2006).

Automatic Elicitation of Aggression by Situational Stimuli

The Weapons Effect

Many of my investigations since the late 1960s were concerned with the role of environmental stimuli closely associated with aggression generally (e.g., Berkowitz, 1964a). [p. 108 ↓] My experiment demonstrating the “weapons effect” (Berkowitz and LePage, 1967) is a good example. Since guns are connected with aggression generally, I had believed the mere sight of a weapon would increase the aggressive inclinations in people disposed to be aggressive at that time. Hostile ideas would be elicited, and aggression-related motor reactions might even be set into operation, so that the viewers would then lash out at their target, especially if they are ready to attack someone at the time and their restraints against aggression are correspondingly weak.

Although there have been several failures to replicate the original Berkowitz – LePage findings, an impressive number of studies support the existence of such a weapons effect, and moreover, these confirming investigations have been conducted in other countries, including Belgium, Croatia,³ Italy, and Sweden, as well as the US (see Berkowitz, 1993). I will not here go into a detailed review of this research, but one point is worth emphasizing. According to several experiments conducted by Turner and his associates (see Berkowitz, 1993), the heightened aggression can be displayed even when the research participants are not aware they are taking part in an experiment. In one of these studies, Turner and his colleagues set up a booth at a college-sponsored carnival and invited students to throw sponges at a target person, allowing them to “assault” the target as often as they wished. More sponges were hurled at the target when a rifle was lying nearby than when no weapon was present.

Effects of Observed Aggression

My associationistic perspective also prompted me to study the effects of seeing aggression on the movie and television screen. Where some proponents of psychoanalytic theorizing had argued that the sight of others fighting would have a cathartic effect, “purging” the viewers’ aggressive impulses, I thought the witnessed violence would function as an aggressive cue automatically evoking associated aggression-related ideas and motor impulses.

My early research confirmed this expectation (see Berkowitz, 1964a, 1964b, 1965, 1993). Geen and O’Neal (1969), following the Wisconsin paradigm, added to this evidence. They demonstrated that if people who had seen a violent movie are physiologically greatly aroused by an irrelevant source soon afterward, they become highly punitive in their subsequent judgments of an available target. Very much in accord with Hullian behavior theory (see Miller, 1959), the “irrelevant arousal” had strengthened the movie-induced aggressive reactions.

Perhaps more interestingly, several Wisconsin studies (e.g., Berkowitz, 1965) showed that cognitive processes could also influence how the observers acted after seeing the violent movie. Under some conditions – such as when the depicted assaults were portrayed as justified – the viewers’ inhibitions against aggression were reduced, and as a consequence, those who had watched this “legitimate” aggression were apt to retaliate severely against someone nearby who had annoyed them earlier. These results, replicated a number of times, have some disturbing implications. In many, perhaps most, violent movies the “good guy” beats up the “villains,” giving them the trashing they supposedly deserve. The hero’s aggression is viewed as justified. This is just the kind of depicted aggression that is especially likely to enhance the audience members’ willingness to assault the “bad persons” in their own lives, at least for a short time afterward.

This cognitive effect obviously can work together with the viewers’ associations in affecting their reactions to movie violence, as one of the Wisconsin studies demonstrates. In this experiment (Berkowitz, 1965), capitalizing on the fact that the university then had a boxing team, each male participant was paired with the

experimenter's confederate posing as a student who was either very interested in college boxing or was a speech major. After a brief exchange with the [p. 109 ↓] confederate, in which this person either angered the participant by making disparaging remarks or treated him in a neutral manner, the participant was given a short synopsis of the prize-fight scene they would view. In this summary, the prize fight loser in the movie was either said to be a nasty person – so that the beating he received would be regarded as justified – or was depicted in a more favorable light – and the aggression was thus unjustified. And also, as was standard in many Wisconsin experiments, immediately after the prize-fight movie each participant then was given an opportunity to administer electric shocks to the target (i.e., the confederate), supposedly as the participant's evaluation of the confederate's solution to an assigned problem.

The angering target was punished most severely when both associative and cognitive influences operated: when the confederate was linked with the aggression on the screen and the witnessed aggression was made to seem justified. The target's association with the observed violence led him to automatically draw stronger attacks and the viewer's favorable interpretation of the aggressive scene reduced their restraints against acting aggressively.

Automatic and Controlled Cognitive Processing

Because of findings such as these, I now couch my analyses in terms of such notions as *automatic processing* and *controlled cognitive processing*, although I continue to devote more attention to the former, automatic reactions governed primarily by associations. Generally speaking, following Schneider and Shiffrin (1977), automatic processing is fast, effortless, obligatory in that attention to a stimulus is sufficient to trigger associated responses, is unconscious, and can act in parallel with other processes. Controlled processing, on the other hand, requires that attention is clearly focused on particular aspects of the situation, that effort has to be exerted, and there is an intention. Reductions in cognitive capacity harm controlled processing but not automatic processing.

My 1984 paper (Berkowitz, 1984) on a cognitive-neoassociation analysis of media effects was a harbinger of this change in my terminology. I noted there that situational occurrences, such as witnessed violence, can prime semantically related thoughts, heightening the chances that the viewers will have other aggressive ideas and even aggressive action tendencies. John Bargh and his colleagues (e.g., Bargh and Williams, 2006; Todorov and Bargh, 2002), perhaps the most prominent of the contemporary researchers investigating automatically elicited social behavior, have advanced a very similar formulation, and have demonstrated in a number of clever ways how social acts can be triggered automatically by particular aspects of the surrounding situation (e.g., Chen and Bargh, 1997).

Automatic Aggressive Reactions to Physically Unattractive Targets

In my earlier discussion of displaced aggression, I proposed that an available target's negative characteristics promote aggressive reactions. Another Wisconsin experiment (Berkowitz and Frodi, 1979; summarized in Berkowitz, 2008), extended this principle to those instances in which the available target had decidedly unattractive physical qualities. In this study, previously annoyed female undergraduates watched a boy, shown on a TV monitor, work on his tasks, and believed their job was to correct his performance over a series of trials. Unbeknownst to the women, they were actually seeing a previously prepared videotape in which the youngster had been made either funny-looking or normal in appearance and, cross-cutting this variation, either stuttered or spoke normally. To enhance the likelihood the participants would respond automatically to what they saw, the women were given another assignment to carry out as they observed the youngster's actions, and were asked to give the child a blast of [p. 110 ↓] noise whenever they thought the boy made a mistake.

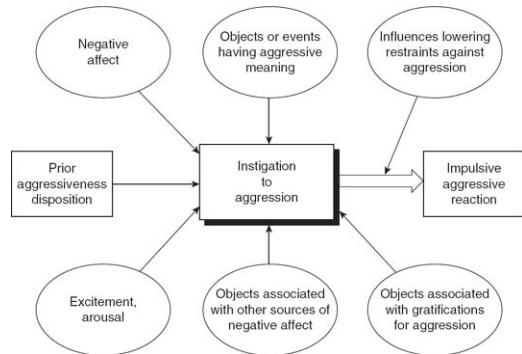
The boy's physical characteristics clearly influenced the punishment the participants gave him for his mistakes (see Figure 2 in Berkowitz, 2008). Evidently still annoyed by the experimenter's earlier treatment of them and somewhat distracted by their own task, they were harsher to the youngster when he stuttered than when he spoke normally,

and also, funny looking rather than normal in appearance. The target's negative characteristics evidently had automatically elicited stronger aggressive reactions.

Cognitions Intensifying Automatically Elicited Aggression

Cognitions can intensify or suppress such automatically activated inclinations. These automatic aggressive reactions can be strengthened, for example, when viewers actively think of the witnessed violence they see as "real" – an actual occurrence – and not "fake" or staged. Many of us are especially apt to associate observed aggression with our own life circumstances when this witnessed occurrence seems real (see Berkowitz and Alioto, 1973: 207). We may then become psychologically quite involved in what happens, so that if we're disposed to be assaultive (and the witnessed aggression appears justified), whatever automatically engendered aggressive inclination we have is strengthened.

Figure 31.2 Influences affecting the magnitude of an impulsive aggressive reaction



People are especially likely to associate the behavior they see with themselves when they identify with the actor, that is, when they actively think of themselves as carrying out the witnessed behavior. Yet another Wisconsin experiment (Turner and Berkowitz, 1972) demonstrated that this identification with the observed aggressor can strengthen the angry viewers' own aggressive inclinations. Each male participant was

first provoked by his partner's unfavorable evaluations of his solutions to an assigned problem, and then three conditions were established before our standard brief fight movie was shown: in one the participant was asked to imagine himself as the fight [p. 111 ↓] winner (who gave the loser a bad beating in the film), in another condition he was to think of himself as the fight judge, and in the last case the participant was not given any imagination role. After this, the participant judged his partner's solutions to the problems he had been given, delivering electric shocks to that man as his evaluations.

As we had expected, the angry people who had imagined themselves as the fight winner delivered much more severe punishment to the person who had offended them than did any of the other groups. In identifying with the movie character who was pummeling his antagonist, they evidently vicariously engaged in aggression themselves, thereby strengthening their aggressive urge.

Cognitions Restraining Aggressive Reactions

Cognitions can also promote the reduction of aggressive inclinations in a number ways – for example, through bringing about a desirable self-regulation (e.g., see Baumeister and Vohs, 2004). I, here, will confine myself to my own research, and will summarize some investigations demonstrating the involvement of cognitive processes in the self-regulation of automatically evoked aggression.

Much of my research in the last few decades of my academic career was focused on the notion that fairly strong displeasure can give rise to an aggressive urge (see Berkowitz, 1993, 2003). In testing this contention, my colleagues and I made our experimental participants experience negative affect in a wide variety of ways, for instance by asking them to engage in unpleasant and stressful physical activities. And very often, in these early studies, as is quite common in investigations of this kind, we asked the participants to rate their feelings before they were given an opportunity to express their hostility. Much to our surprise, and quite contrary to the “demand characteristics” notion, in some of our initial studies the negative affect led to a *low* level of expressed hostility. It then occurred to me that in rating their feelings the participants

could have become highly aware of their unpleasant affect and hostile tendencies. They might have then regarded these emotional reactions as inappropriate at that time and consequently sought to regulate their expressed reactions.

This interpretation is very similar to the reasoning advanced by Carver and Scheier (1981) in their theorizing about self-awareness and self-regulation. These writers held, in essence, that when people become highly aware of themselves, they also become very conscious of their personal standards pertinent to the kind of situation they are in. If there is a discrepancy between these standards and what they are inclined to do at that moment, they will then attempt to minimize this discrepancy and act in accord with their personal values. Our instructing the students to describe their feelings after the experimental manipulation could have had a similar effect: they might have become very aware of themselves as well as their hostile inclinations. Believing it was improper to be nasty to their experimental partner in the current situation, they presumably then avoided expressing anything very negative about this person.

This evidently happened in an experiment by Berkowitz and Troccoli (1990). Half of the female participants were individually made physically very uncomfortable by requiring them to hold their nondominant arm outward, whereas the other women only rested their arm on the desk before them. Then, while maintaining their arm in the stipulated position, the participants heard a tape recording on which a supposed job applicant talked about herself. When this ended half of the women were asked to think about their feelings at that time, whereas their counterparts were given a distracting word-association task. Finally, about five minutes after they had started, and with the participants still maintaining the specified arm position, all of them evaluated the female job applicant.

Not surprisingly, the women whose attention had been diverted from themselves showed the usual feelings-congruence effect; [p. 112 ↓] those made physically uncomfortable were harsher in their judgments of the applicant than their more comfortable counterparts. By contrast, however, the people who had attended to their negative feelings apparently showed the self-awareness induced self-regulation; the uncomfortable participants actually were kinder to the job applicant than their more comfortable peers.

Yet another of our studies suggests that it is the actions produced by the negative affect that is regulated, not the feeling itself. In this experiment (Berkowitz and Jo, 1992, unpublished) the female participants' feelings were manipulated in such a way that they were unlikely to know the source of whatever affect they experienced. Following the procedure described by Strack et al. (1988), they were asked to hold their mouths in a particular way so that they adopted either a smiling or frowning facial expression. Shortly after this expression was established some of the women rated their feelings and thus became highly aware of their affective state, whereas the other participants were distracted by having them list several word associates. All of the participants then read an autobiographical statement supposedly written by a job applicant, and then, as in the Berkowitz and Troccoli (1990) experiment, evaluated her.

The results obtained here parallel the findings in the previous investigation. Here again, the judgments were congruent with the participants' feelings when their attention had been diverted from themselves; the smiling women assigned fewer bad qualities to the target than did their frowning counterparts. And also, as had been found before, the participants' attention to their feelings led to the feeling-judgment incongruence; the frowning women were less harsh in their judgments than their smiling peers. The active cognitive processing produced by the self-directed attention evidently had caused these latter persons to suppress their display of hostility to the job applicant.

Impulsive Aggression

The results summarized here, together with other research findings, such as those obtained by Bargh and his colleagues (e.g., Bargh and Williams, 2006; Todorov and Bargh, 2002), demonstrate that aggressive verbal and motor actions can be automatically elicited by aggression-related stimuli. I have often (e.g., Berkowitz, 1993, 2003) referred to these attacks as *impulsive* in the sense that they are carried out with little thought in response to situational features, and proceed with little, if any, conscious guidance and attention. Some instances of impulsive aggression are largely the result of overly rapid decision-making, but many impulsive assaults are often primarily a product of factors affecting the disinhibition of situationally induced reactions, with little part played by "higher order" cognitive processes.

Bushman and Anderson (2001) have questioned this contention that many aggressive actions are impulsive attacks automatically activated by relevant situational stimulation rather than because of a decision that had been reached. They maintained, instead, that it is advisable to think of every aggressive action as the product of a decision process greatly affected by the individual's knowledge structures. Nonetheless, in my view, much of the research into individual differences in aggressive behavior support my distinction between impulsive, automatically elicited assaults, and cognitively controlled, chosen attacks. Thus, where Dodge (e.g., Crick and Dodge, 1996), in differentiating between *reactive* and *proactive* aggression, has emphasized the role of information processing in both of these cases (also see Berkowitz, 2008), other investigations (e.g., Raine et al., 2006) have noted that reactive aggressors are often apt to be highly impulsive. For that matter, Barratt (1999), well known for his studies of impulsivity, seemed to prefer the concepts *impulsive* versus *premeditated* aggressive acts instead of Dodge's labels.

[p. 113 ↓]

Of course, a good many factors can affect the magnitude of an impulsive attack on an available target, and [Figure 31.2](#) lists the influences suggested by my CNA model. Up to now in the present discussion I have dealt mostly with the roles of negative affect and the priming established by objects or events having an aggressive meaning, but here I would like to say a bit more about the weakness of restraints against aggression in impulsivity.

Impulsivity has been operationally defined in a great many different ways (e.g., see Dickman, 1990; White et al., 1994), but according to several investigations, impulsive actions often vary along two underlying dimensions: one dealing mostly with unrestrained acts and the other reflecting an incomplete and inadequate information processing so that overly quick decisions are reached. In the earlier of these studies, Dickman (1990) uncovered two factors that appear to tap these two dimensions. His items (all self-reported) with the highest loadings on Factor 1 included, "I don't like to make decisions quickly, even simple decisions" (disagree), and "I am uncomfortable when I have to make up my mind rapidly," (disagree), suggesting the factor has to do with very hasty decision-making. Factor 2, on the other hand, based on items such as, "I often say and do things without considering the consequences," and "I often get into

trouble because I don't think before I act," appears to reflect a tendency to impetuous actions. These two factors had a low but significant positive correlation. Another study (Endicott et al., 2006), in investigating the correlates of self-reported impulsivity, also differentiated between a "nonplanning" impulsiveness and a "motor impulsiveness," with the latter having to do primarily with "inhibitory dyscontrol." Yet another investigation also uncovered these two types of impulsivity. White and Moffitt and their colleagues (1994) factor analyzed the relationships among the 11 different impulsivity measures they employed in their study. Two correlated but distinct forms of impulsivity were identified. The first of these, termed "behavioral impulsivity," appeared to reflect a lack of behavioral control in that "the variables with the highest loading on this factor were those that tapped disinhibited, undercontrolled behavior." The second factor, labeled "cognitive impulsivity," had more to do with a deficiency in "effortful and planful cognitive performance." Both kinds of impulsivity were associated with delinquency in the White et al. (1998) research, although the data indicated that the relationship between behavioral impulsivity and delinquency was independent of intelligence.

Conclusion

The aggressive reactions I and others have seen in the laboratory experiments cited here admittedly have been quite weak. But nevertheless, I argue that the findings obtained in these studies, especially when taken together with more naturalistic investigations, point to a possible explanation of many instances of domestic and even criminal violence. In these assaults (as [Figure 31.2](#) suggests) people disposed to be aggressive, because of their personalities and/or the strong negative affect they were experiencing, could well have impulsively attacked an available target, especially one having (for the aggressor) decidedly negative qualities – without necessarily having consciously decided to do so. Increasing numbers of legal scholars and philosophers now recognize the problems that are raised for legal justice systems by psychological findings such as this. Law codes typically use judgments of an offender's intentions in deciding what punishment he or she should receive for a crime. But what if there was little if any conscious intention? In the case of, say, domestic violence, shouldn't the external influences automatically eliciting aggressive inclinations be considered at least as mitigating circumstances?

[p. 114 ↓]

One answer, of course, is that people should restrain their aggressive urges, and the threat of punishment could promote this necessary self-regulation. But there are many more, and probably better, ways of getting people to control themselves, as the burgeoning literature on self-regulation (e.g., Baumeister and Vohs, 2004) shows. The previously mentioned experiment by Berkowitz and Troccoli (1990) indicates, for example, along with the Carver and Scheier (1981) analysis, that persons disposed to be assaultive would benefit from both the learning of nonaggressive standards of conduct and the heightening of their self-awareness when they are aggressively inclined. Clearly, further research into the self-regulation of aggressive impulses is highly desirable.

In addition to emphasizing the role of aggression-eliciting situational stimuli, CNA also gives special attention to aversive occurrences generally and decidedly negative affect in particular. Experienced stress is just one variation on this theme. Studies mentioned earlier have documented how people under stress have become assaultive. Combat-induced stress is another example. As a case in point, the New York Times (January 2, 2009) examined the cases of over 120 veterans of the Iraq or Afghanistan wars who were charged with homicide after their return. The reporter concluded that combat trauma and stress “appeared to have set the stage for [many of] the crimes” (p. A12). Interestingly, this article also noted that a number of the accused killers had previously tried to commit suicide, supporting my suggestion (Berkowitz, 1993) that intense depression can also generate an aggressive urge, directed against others as well as to the self. And here too, much more has to be learned. Although I have repeatedly held that strong negative affect generates an aggressive inclination (which may or may not be expressed openly; see [Figure 31.1](#)), it is intuitively likely that some kinds of intense unpleasant feelings are more aggression-inducing than other kinds. My guess here is that an agitated displeasure is more apt to have assaultive consequences than, say, a “flat,” listless mood. We don’t know, however, whether this is indeed so, or even what just factors produce these different negative feelings.

All in all, then, the CNA model spelled out in this paper is just a preliminary formulation, one that leaves many important questions unanswered. But at least it has the virtue

of raising these questions. Hopefully, other investigators will undertake the task of pursuing these problems further.

Notes

1 These discussions can be found in the Vol. 41(4) issue of *Psychological Review* (1941), and in Newcomb and Hartley (eds) (1947) *Readings in Social Psychology*. New York: Holt.

2 The studies by Pedersen, Miller, and their associates (e.g., Pedersen et al., 2000) into “triggered aggression” are an interesting and important extension of the more usual investigations of displaced aggression just summarized. I had intended to cite some of their experiments in this report, but space limitations unfortunately kept me from doing so.

3 In the Croatian study, as an example, the investigator, Miomir Zuzul (see Berkowitz, 1993), showed that previously frustrated schoolchildren were more aggressive toward their peers in a free play situation after they had been exposed to either real or toy guns than after seeing a neutral object, but primarily if the supervising adult had expressed either a neutral or favorable attitude toward aggression generally.

References

Anderson, C. and Anderson, K. Violent crime rate studies in philosophical context: A destructive testing approach to heat and southern culture of violence effects . Journal of Personality and Social Psychology , 70, 740–756. (1996) <http://dx.doi.org/10.1037/0022-3514.70.4.740>

Anderson, C. and Anderson, K. (1998) Temperature and aggression: Paradox, controversy, and a (fairly) clear picture . In R. Geen, ed. and E. Donnerstein (eds), Human Aggression: Theories, Research, and Implications for Social Policy , pp. 247–298. San Diego: Academic Press. <http://dx.doi.org/10.1016/B978-012278805-5/50011-0>

Anderson, C., Deuser, W. and DeNeve, K. Hot temperatures, hostile affect, hostile cognition, and arousal: Tests of a general model of affective aggression . *Personality and Social Psychology Bulletin* , 21, 434–448. (1995) <http://dx.doi.org/10.1177/0146167295215002>

Bandura, A. and Walters, R. (1963) *Social Learning and Personality Development* . New York: Holt, Rinehart and Winston.

Bargh, J. and Williams, E. The automaticity of social life . *Current Directions in Psychological Science* , 15, 1–4. (2006) <http://dx.doi.org/10.1111/j.0963-7214.2006.00395.x>

Baron, R., Byrne, D. and Griffitt, W. (1974) *Social Psychology* . Boston: Allyn and Bacon.

Barratt, E. Impulsive and premeditated aggression: A factor analysis of self-reported acts . *Psychiatry Research* , 86, 163–173. (1999) <http://dx.doi.org/10.1016/S0165-1781%2899%2900024-4>

Baumeister, R. and Vohs, K. (2004) *Handbook of Self-regulation* . New York: Guilford Press.

Berkowitz, L. The expression and reduction of hostility . *Psychological Bulletin* , 55, 257–283. (1958) <http://dx.doi.org/10.1037/h0045629>

Berkowitz, L. Anti-Semitism and the displacement of aggression . *Journal of Abnormal and Social Psychology* , 59, 182–187. (1959) <http://dx.doi.org/10.1037/h0048227>

Berkowitz, L. (1963) *Aggression: A Social Psychological Analysis* . New York: McGraw-Hill.

Berkowitz, L. Aggressive cues in aggressive behavior and hostility catharsis . *Psychological Review* , 71, 104–122. (1964a) <http://dx.doi.org/10.1037/h0043520>

Berkowitz, L. The effects of observing violence . *Scientific American* , 210, 35–41. (1964b) <http://dx.doi.org/10.1038/scientificamerican0264-35>

Berkowitz, L. Some aspects of observed aggression . Journal of Personality and Social Psychology , 2, 359–369. (1965) <http://dx.doi.org/10.1037/h0022221>

Berkowitz, L. (ed.) (1969) Roots of Aggression: A Re-examination of the Frustration-aggression Hypothesis . New York: Atherton Press.

Berkowitz, L. Aversively stimulated aggression: Some parallels and differences in research with animals and humans . American Psychologist , 38, 1135–1144. (1983) <http://dx.doi.org/10.1037/0003-066X.38.11.1135>

Berkowitz, L. Some effects of thoughts on anti- and prosocial influences of media events: A cognitive-neoassociation analysis . Psychological Bulletin , 95, 410–427. (1984) <http://dx.doi.org/10.1037/0033-2909.95.3.410>

Berkowitz, L. The frustration-aggression hypothesis: Examination and reformulation . Psychological Bulletin , 106, 59–73. (1989) <http://dx.doi.org/10.1037/0033-2909.106.1.59>

Berkowitz, L. (1993) Aggression: Its Causes, Consequences, and Control . New York: McGraw-Hill.

Berkowitz, L. (2003) Affect, aggression, and antisocial behavior . In R. Davidson, ed. , K. Scherer, ed. and H. Goldsmith (eds), Handbook of Affective Sciences , pp. 804–823. Oxford: Oxford University Press.

Berkowitz, L. On the consideration of automatic as well as controlled psychological processes in aggression . Aggressive Behavior , 34, 117–129. (2008) <http://dx.doi.org/10.1002/ab.20244>

Berkowitz, L. and Alioto, J. The meaning of an observed event as a determinant of its aggressive consequences . Journal of Personality and Social Psychology , 28, 206–217. (1973) <http://dx.doi.org/10.1037/h0035730>

Berkowitz, L., Cochran, S. and Embree, M. Physical pain and the goal of aversively stimulated aggression . Journal of Personality and Social Psychology , 40, 687–700. (1981) <http://dx.doi.org/10.1037/0022-3514.40.4.687>

Berkowitz, L. and Frodi, A. Reactions to a child's mistakes as affected by her/his looks and speech . *Social Psychology Quarterly* , 42, 420–425. (1979) <http://dx.doi.org/10.2307/3033814>

Berkowitz, L. and Holmes, D. A further investigation of hostility generalization to disliked objects . *Journal of Personality* , 28, 427–442. (1960) <http://dx.doi.org/10.1111/j.1467-6494.1960.tb01630.x>

Berkowitz, L. and Knurek, D. Label-mediated hostility generalization . *Journal of Personality and Social Psychology* , 13, 200–206. (1969) <http://dx.doi.org/10.1037/h0028125>

Berkowitz, L. and LePage, A. Weapons as aggression-eliciting stimuli . *Journal of Personality and Social Psychology* , 7, 202–207. (1967) <http://dx.doi.org/10.1037/h0025008>

Berkowitz, L. and Troccoli, B. Feelings, direction of attention, and expressed evaluations of others . *Cognition and Emotion* , 4, 305–325. (1990) <http://dx.doi.org/10.1080/02699939008408080>

Bower, G. Mood and memory . *American Psychologist* , 36, 129–148. (1981) <http://dx.doi.org/10.1037/0003-066X.36.2.129>

Brewer, M. and Brown, R. (1998) Intergroup relations . In D. Gilbert, ed. , S. Fiske, ed. , and G. Lindzey (eds) *Handbook of Social Psychology*, Vol. 2 , 4th Edition, pp. 554–594. New York: McGraw-Hill.

Bushman, B. and Anderson, C. Is it time to pull the plug on the hostile versus instrumental aggression dichotomy? *Psychological Review* , 108, 273–279. (2001) <http://dx.doi.org/10.1037/0033-295X.108.1.273>

Carver, C. and Scheier, M. (1981) *Attention and Self-regulation* . New York: Springer. <http://dx.doi.org/10.1007/978-1-4612-5887-2>

Chen, M. and Bargh, J. Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation . *Journal of Experimental Social Psychology* , 33, 541–560. (1997) <http://dx.doi.org/10.1006/jesp.1997.1329>

Cohn, E. and Rotton, J. Assault as a function of time and temperature: A moderator-variable time-series analysis . *Journal of Personality and Social Psychology* , 72, 1322–1334. (1997) <http://dx.doi.org/10.1037/0022-3514.72.6.1322>

Crick, N. and Dodge, K. Social information-processing mechanisms in reactive and proactive aggression . *Child Development* , 67, 993–1002. (1996) <http://dx.doi.org/10.2307/1131875>

Dickman, S. Functional and dysfunctional impulsivity: Personality and cognitive correlates . *Journal of Personality and Social Psychology* , 58, 95–102. (1990) <http://dx.doi.org/10.1037/0022-3514.58.1.95>

Dollard, J., Doob, L., Miller, N., Mowrer, O. and Sears, R. (1939) *Frustration and Aggression* . New Haven: Yale University Press. <http://dx.doi.org/10.1037/10022-000>

Duclos, S., Laird, J., Schneider, E., Sexton, M., Stern, L. and Van Lighten, O. Emotion-specific effects of facial expressions and postures on emotional experience . *Journal of Personality and Social Psychology* , 57, 100–108. (1989) <http://dx.doi.org/10.1037/0022-3514.57.1.100>

Endicott, P., Ogloff, J. and Bradshaw, J. *Personality and Individual Differences* , 41, 285–294. (2006) <http://dx.doi.org/10.1016/j.paid.2006.01.011>

Fernandez, E. and Turk, D. The scope and significance of anger in the experience of chronic pain . *Pain* , 61, 165–175. (1995) <http://dx.doi.org/10.1016/0304-3959%2895%2900192-U>

Fernandez, E. and Wasan, A. (in press) The anger of pain sufferers: Attributions to agents and appraisals of wrongdoing . In M. Potegal, ed. , G. Stemmler, ed. , and C. Spielberger (eds), *Handbook of Anger* . New York: Springer.

Fitz, D. A renewed look at Miller's conflict theory of aggression displacement . Journal of Personality and Social Psychology , 33, 725–732. (1976) <http://dx.doi.org/10.1037/0022-3514.33.6.725>

Geen, R. Effects of frustration, attack, and prior training in aggressiveness upon aggressive behavior . Journal of Personality and Social Psychology , 9, 316–321. (1968) <http://dx.doi.org/10.1037/h0026054>

Geen, R. and O'Neal, E. Activation of cue-elicited aggression by general arousal . Journal of Personality and Social Psychology , 11, 289–292. (1969) <http://dx.doi.org/10.1037/h0026885>

Green, D., Glaser, J. and Rich, A. From lynching to gay bashing: The elusive connection between economic conditions and hate crime . Journal of Personality and Social Psychology , 75, 82–92. (1998) <http://dx.doi.org/10.1037/0022-3514.75.1.82>

GreenwoodK., Thurston, R., Rumble, M., Waters, S. and Keefe, F. Anger and persistent pain: Current status and future directions . Pain , 103, 1–5. (2003) <http://dx.doi.org/10.1016/S0304-3959%2803%2900132-5>

Hewitt, L. (1974) Who will be the target of displaced aggression? In J. De Wit, ed. and W. Hartup (eds), Determinants and Origins of Aggressive Behavior , pp. 217–223. The Hague: Mouton.

Huesmann, L. An information processing model for the development of aggression . Aggressive Behavior , 14, 13–24. (1988) <http://dx.doi.org/10.1002/1098-2337%281988%2914:1%3C13::AID-AB2480140104%3E3.0.CO;2-J>

Keltner, D., Ellsworth, P. and Edwards, K. Beyond simple pessimism: Effects of sadness and anger on social perception . Journal of Personality and Social Psychology , 64, 740–752. (1993) <http://dx.doi.org/10.1037/0022-3514.64.5.740>

Lewis, M. (1993) The development of anger and rage . In R. Glick, ed. and S. Roose (eds), Rage, Power, and Aggression , pp. 148–168. New Haven: Yale University Press.

Loew, C. Acquisition of hostile attitude and its relation to aggressive behavior . Journal of Personality and Social Psychology , 5, 335–341. (1967) <http://dx.doi.org/10.1037/h0024259>

Marcus-Newhall, A., Pedersen, W., Carlson, M. and Miller, N. Displaced aggression is alive and well: A meta-analytic review . Journal of Personality and Social Psychology , 78, 670–689. (2000) <http://dx.doi.org/10.1037/0022-3514.78.4.670>

MacDonald, G. and Leary, M. Why does social exclusion hurt? The relation between social and physical pain . Psychological Bulletin , 131, 202–223. (2005) <http://dx.doi.org/10.1037/0033-2909.131.2.202>

Miller, N. The frustration-aggression hypothesis . Psychological Review , 48, 337–342. (1941) <http://dx.doi.org/10.1037/h0055861>

Miller, N. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization . Journal of Abnormal and Social Psychology , 43, 155–178. (1948) <http://dx.doi.org/10.1037/h0056728>

Miller, N. (1959). Liberalization of basic S-R concepts: Extensions to conflict behavior, motivation and social learning . In S. Koch (ed.), Psychology: A Study of a Science , 2, 196–292. New York: McGraw-Hill.

Nisbett, R. and Cohen, D. (1996) Culture of Honor: The Psychology of Violence in the South . Boulder, CO: Westview Press.

Parrott, D., Zeichner, A. and Hoover, R. Sexual prejudice and anger network activation: Mediating role of negative affect . Aggressive Behavior , 32, 7–16. (2006) <http://dx.doi.org/10.1002/ab.20101>

Passman, R. and Mulhern, R. Maternal punitiveness as affected by situational stress: An experimental analogue of child abuse . Journal of Abnormal Psychology , 86, 565–569. (1977) <http://dx.doi.org/10.1037/0021-843X.86.5.565>

Pedersen, W., Gonzales, C. and Miller, N. The moderating effect of trivial triggering provocation on displaced aggression . *Journal of Personality and Social Psychology* , 78, 913–927. (2000) <http://dx.doi.org/10.1037/0022-3514.78.5.913>

Quigley, B. and Tedeschi, J. Mediating effects of blame attributions on feelings of anger . *Personality and Social Psychology Bulletin* , 22, 1280–1288. (1996) <http://dx.doi.org/10.1177/01461672962212008>

Raine, A., Dodge, K., Loeber, R., Gatzke-Kopp, L., Lynam, D., Chandra, R., Stouthamer-Loeber, M. and Lie, J. The reactive-proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys . *Aggressive Behavior* , 32, 159–171. (2006) <http://dx.doi.org/10.1002/ab.20115>

Schneider, W. and Shiffrin, R. Controlled and automatic human information processing: I. Detection, search, and attention . *Psychological Review* , 84, 1–66. (1977) <http://dx.doi.org/10.1037/0033-295X.84.1.1>

Stifter, C. and Grant, W. Infant responses to frustration: Individual differences in the expression of negative affect . *Journal of Nonverbal Behavior* , 17, 187–204. (1993) <http://dx.doi.org/10.1007/BF00986119>

StrackF., Martin, L. and Stepper, S. Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis . *Journal of Personality and Social Psychology* , 54, 768–777. (1988) <http://dx.doi.org/10.1037/0022-3514.54.5.768>

Straus, M. (1980) Stress and child abuse . In H. Kempe, ed. and R. Helfer (eds), *Stress and Child Abuse* , 3rd Edition. Chicago: University of Chicago Press.

Todorov, A. and Bargh, J. Automatic sources of aggression . *Aggression and Violent Behavior* , 7, 53–68. (2002) <http://dx.doi.org/10.1016/S1359-1789%2800%2900036-7>

Turner, C. and Berkowitz, L. Identification with film aggressor (covert role taking) and reactions to film violence . *Journal of Personality and Social Psychology* , 21, 256–264. (1972) <http://dx.doi.org/10.1037/h0032267>

White, J., Moffitt, T., Caspi, A., Bartusch, D., Needles, D. and Stouthamer-Loeber, M. Measuring impulsivity and exploring the relationship to delinquency . Journal of Abnormal Psychology , 103, 192–205. (1994) <http://dx.doi.org/10.1037/0021-843X.103.2.192>

<http://dx.doi.org/10.4135/9781446249222.n31>

Handbook of Theories of Social Psychology

Need-to-Belong Theory

Contributors: Roy F. Baumeister

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Need-to-Belong Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n32>

Print pages: 121-141

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n32>

[p. 121 ↓]

Chapter 32: Need-to-Belong Theory

Roy F.Baumeister

Abstract

The motivation to form and sustain at least a minimum amount of social connections is one of the most powerful, universal, and influential human drives. It shapes emotion, cognition, and behavior. It explains self-esteem as an internal measure of one's chances of having good relationships. Different ways of satisfying the need to belong can explain gender differences in personality and roles and even reinterpret the history of gender politics, on the assumption that women emphasize close, intimate relationships whereas men are oriented toward larger networks of shallower relationships. Studies of rejection show that thwarting the need to belong produces drastic and sometimes puzzling effects, including increases in aggression and self-destructive acts, and decreases in helpfulness, cooperation, self-control, and intelligent thought. Lab rejection studies produce an emotional numbness that has pointed the way to question the basic functions of emotion and how emotion affects behavior. Culture, which is the ultimate achievement and form of human social life, depends on belongingness, and that observation offers a powerful basis for understanding human nature and many distinctively human traits.

Introduction

People have a basic need to belong. They are motivated to form and maintain social relationships. This deceptively simple idea was the topic of a review article by Mark Leary and me that kept me busy for several years in the early 1990s, up until its publication (Baumeister and Leary, 1995). It changed my thinking permanently and has formed the basis for much of my subsequent research and theorizing. That article has

been cited over a thousand times in the research literature, and so perhaps it has had influence on other people as well.

The idea itself is hardly revolutionary or controversial. Of course people like to be connected to other people, by and large. Yet this simple idea led in many unexpected directions. Its influence on cognition, emotion, and behavior is extensive. Moreover, it raises important basic questions about human nature, culture, gender, emotion, and how the human psyche functions.

How the Project Began

The crucial steps that led Leary and me to write the “need-to-belong” paper were taken in a series of conversations and exchanges in the late 1980s. Inevitably, however, one can look back and point to earlier factors that [p. 122 ↓] were relevant. In my own case, this includes a spirit of contrarian inquiry that has been a theme of my life and has contributed to many of my career’s biggest problems and failures as well as my best successes and achievements. My (admittedly suspect) autobiographical recollection of the roots of this go back to some point in junior high school, back when I was first assigned to write papers for school. My mother taught high school and pointed out that a teacher who grades a stack of papers on an assigned topic quickly gets bored when one after another says pretty much the same thing, even if that common theme is essentially sound and correct. This occurred long before grade inflation and during the era of heavy tracking, so even in a public school the competition for top grades was intense. I realized that my chances of getting an A would be better if I could say something different from the bulk of other students’ papers. I learned to think about a problem not just in terms of question and answer, or problem and solution, but also in terms of what the dominant approach might be overlooking. I wanted my ideas to be not just correct but also unusual.

The same contrarian approach can serve one well in science. If you discern what the prevailing trends and biases are, you can guess what is likely to be missed. I have followed this strategy throughout my career, by habit and now by inclination. It turns out to have risks as well as advantages. When you look where no one else is looking, you find more interesting things more easily than when you hunt in the same

hunting grounds that everyone else uses. People think you're creative, because your work is different from the mainstream. But they also value it less (it is, after all, less relevant to what everyone else is working on). It is also much more likely to stray into areas deemed politically incorrect: Political correctness defines what thoughts are not supposed to be considered, and so someone who looks at what others ignore easily strays into danger of being politically incorrect, a danger that is not to be taken lightly.

My university education took place in the 1970s, when psychology was entering into what later became known as the cognitive revolution. Attribution theory was already flourishing in social psychology. The study of attitudes, especially cognitive dissonance, was at its peak as well. Both these approaches had social psychologists thinking heavily about what happens inside the person. As a contrarian, therefore, I attended to what went on between persons. Sure enough, I felt, interpersonal processes were unappreciated and understudied, and so I found a good start there. For example, where others spoke about self-esteem, I focused on self-presentation – not the self's concept of itself, but its concern with how others saw it. Indeed, in the mid 1980s, my first edited book was called *Public Self and Private Self*, in which I asked contributors to reflect on how these inner and outer processes were related.

One of the contributions to that volume was an unusual chapter by Greenberg et al. (1986). I had read some laboratory studies by two of them focusing on how interpersonal events impacted private self-esteem (Greenberg and Pyszczynski, 1985) and invited them on that basis, but instead of writing about that work, they developed a grand theory about human motivation based loosely on some anthropological writings (Becker, 1973). The chapter in my book was in fact the first publication on terror management theory, which has gone on to become quite influential. Based on this early connection, those fellows and I sought each other out at some conferences.

During the 1980s, Bibb Latané held small conferences every summer at a property he owned at Nags Head, North Carolina, on the beach. There was usually one conference on the self, and often the terror management fellows and I all attended. All three of them would be present, and so their work came to constitute a major theme those conferences. They were deliberately controversial.

A central point of their theory was that anxiety was mainly caused by fear of death. (This is the “terror” in the theory’s name.) [p. 123 ↓] To be sure, the evidence for terror and anxiety was never strong, and more recent formulations have downplayed the terror aspect, but initially it was central. The core idea was thus highly individualistic: A person knows that he or she is going to die, contemplates mortality, has anxiety as a result, and therefore performs a variety of behaviors to escape from this anxiety. Other people are not directly or necessarily involved in the central drama and theme of life.

One year Leary and I and Dianne Tice were all in attendance and listened to the terror management presentations. We had by now heard them multiple times. We discussed the terror management theory of anxiety. We had various reasons for wondering whether anxiety was really based on death. We did not know what anxiety researchers thought, apart from Freud’s early theory that anxiety in men comes from fear of castration, and from his later theory (elaborated by Bowlby and others) that social separation is the root of anxiety.

By now I had a long habit of deliberately considering interpersonal perspectives on things, so the idea of social separation as a theory of what causes anxiety intrigued me. I thought terror management theory could be made much more broadly accurate and appealing if they expanded it to include not just terror of death but also terror of social rejection. I went up to the room where the terror management fellows were staying. Pyszczynski and Solomon were in there, talking about ideas while smoking and drinking in a jovial spirit. I ran the idea by them of expanding their theory to include social exclusion and separation anxiety. They nodded, pondered, and said they liked this idea. They wanted to run it by Greenberg. The next day, however, they told me that Greenberg had vetoed the expansion: death was to be the only source of terror in their theory. I have long wondered how things would have been different had they accepted my overtures, as two of the three of them seemed inclined to do.

Our curiosity piqued and undaunted, Tice and I embarked on a literature review to learn about anxiety. The goal was just to find out what actually causes anxiety. We did find some evidence to support the terror management view: some people do have anxiety over the possibility of injury or death. But by far the biggest stimulus was fear of being rejected and abandoned by others. This evolved into a review article called “Anxiety and social exclusion” (Baumeister and Tice, 1990).

We were not sure where to publish this manuscript. C.R. Snyder had attended some of the Nags Head conferences with us and suggested we publish it in the journal he edited, the *Journal of Social and Clinical Psychology*. He said he would treat it as a target article and get commentaries from various experts. Leary wrote a commentary, as did several others, which turned into quite an intriguing exchange, marred only by an extremely hostile and defensive commentary by the terror management group. Not only were they uninterested in incorporating separation anxiety into their own theory: they rejected the very idea that anxiety could come from fear of social separation.

When this issue was finally published, I thought the episode completed, but in a conversation with Leary the next year he made a crucial comment. He said he thought that our account of anxiety as often stimulated by fear of social exclusion, though largely correct, suffered from not going far enough. He thought the concern over social belongingness versus rejection went beyond just anxiety. His comment was the crucial forward leap that led to the development of our theory.

Assembling the evidence and writing the article on “the need to belong” (Baumeister and Leary, 1995) was a powerful experience. We did not know that this would become the most-cited article that either of us had ever published. We worried at times that we were just trying to prove something that everyone already knows, which is that people are motivated to form social bonds.

Yet the extent of the evidence and the widely varying forms it took continued to [p. 124 ↓] surprise us. We kept coming across new findings and patterns that showed yet another direction where the need to belong applied. We found echoes of the need to belong in unexpected places. People form social bonds remarkably easily and readily, even just based on having experienced electric shock together in an experiment (Latané et al., 1966). They are irrationally reluctant to let bonds break, even continuing to exchange Christmas cards with people they do not know (Kunz and Woolcott, 1976). Many cognitive and emotional patterns are linked with interpersonal connections.

Building the Basic Theory

The core idea of the need to belong is that people have a fundamental, strong, and pervasive motivation to form and maintain at least a certain minimum number of social relationships. The drawbacks of this idea included its simplicity as well as its lack of novelty. Other theorists had said that people have such a drive to form social bonds. In fact, as reviewers would soon point out, this was not an idea that anyone would vehemently disagree with. Most personality theories had at some point said that people have some sort of drive to affiliate with others.

We had therefore hardly discovered a new motivation. Rather, our contribution was to document the power and extent of this motivation. True, most personality theories had acknowledged such a drive, but it was usually in the background, treated as a relatively minor and uninteresting aspect of human nature. Freud, for example, had gradually acknowledged separation anxiety as genuine, possibly stemming from the baby's dependency on its mother, but the centerpiece motivations in his theory were the sexual and aggressive drives.

The phrase "the need to belong" appealed to Leary and me because it conveyed a greater sense of power and urgency than something like an "affiliation motive." However, the term "need" itself was potentially problematic. Is the need to belong truly a need, or just a want? We had to take a detour to explore metatheory: What makes a drive qualify as a need? Our answers, which were developed in the early part of the article, focused on whether actual harm would come from failing to satisfy the motivation. Not getting something you *need* means more and is worse than not getting something you merely *want*. This does not necessarily predict the strength of the motivation, we realized. Sexual desire may be an extremely strong motivation, especially in some people, but ultimately it is a want rather than a need. No one dies or even gets sick from lack of sex. In fact there does not seem to be any notable harm that stems from not having sex, and some people live long, healthy, happy lives without ever having sex.

Now that we had a theory about what qualifies as a need, we had to face the question of whether our so-called need to belong would actually meet those criteria. Fortunately,

the evidence was there. In fact, as we had repeatedly found while reviewing various aspects of the need to belong, the evidence was far stronger and more abundant than we had anticipated. It was not just a matter of a few potential types of harm that could be traced to severe loneliness. Rather, it seemed that the majority of mental and physical illnesses were affected in some way by a lack of social ties. For example, a review of medical research came to the shockingly broad conclusion that mortality rates for *all* major causes of death were higher among people who lacked social bonds than for people who were well connected to other people (Lynch, 1979).

Buoyed by such findings, we felt it justified in asserting that the drive to connect with others was more than a mere desire. What had been up till this point our working title, "the need to belong," became official.

As usual, the paper and its ideas benefited from the journal review process. The reviewers and the journal editor, Robert Sternberg, [p. 125 ↓] pressed us to say just exactly what is meant by belongingness. What specifically is needed in the "need to belong"? Is it just a herd instinct, a craving for affiliation, companionship? Leary and I grappled with this, searched for new evidence, and formulated and revised various ideas. We settled on two central aspects of the need to belong. First, the person wants a framework of mutual concern and caring that extends into past and future. Second, the person wants a series of non-negative interactions. Getting either without the other should amount to half a loaf: partly but not wholly satisfactory.

The framework of mutual concern extending into past and future meant that momentary interactions would not be sufficient, no matter how intimate. The need to belong is not satisfied by any amount of social interaction without the bond. The data went along nicely with this. Toll takers on the turnpike may interact with hundreds, possibly thousands of people every day, but these very brief conversations do not satisfy the need to belong. Even prostitutes, who may have a highly intimate ("lovemaking") interaction with several dozen persons a day, do not find these satisfactory, though perhaps they are spared the loneliness of forest rangers and night watchmen (McLeod, 1982). We found evidence that some prostitutes sacrifice the best chances for making the most money in order to cultivate relationships with repeat customers (Symanski, 1980). Meanwhile, though, even without social contact, people do get some benefit from these bonds. Data on long-distance relationships showed that these often have great

importance, even when there is no contact for long periods of time (Gerstel and Gross, 1982, 1984). For example, studies of married sailors, especially those in submarines, showed that the sailors and the spouses valued the connection even though six months might go by with no contact at all or at best just an occasional letter (Harrison and Connors, 1984). Thus, a bond without frequent interaction was indeed half a loaf: much more than nothing but far less than fully satisfying.

Meanwhile, what about the interactions? The long-distance relationships data indicated relatedness without interaction was missing something, so the interactions do seem to matter. Here, however, we were faced with some contrary findings that prevented us from simply emphasizing face-to-face contact as a second key. The data were strongly suggestive that when face-to-face contacts are hostile, abusive, or demoralizing, they did not produce good results. We then thought that we should stipulate that people needed the interactions to be positive and pleasant, but reviewers objected, quite reasonably we thought, that many old friends and long-married couples seem to find being together satisfying without a great deal of exuberant positivity. The dual-career couples working apart reported that they wished just to sit and watch television with each other, which is hardly an intimate or emotionally intense experience. These arguments pushed Leary and me to accept, at least tentatively, that interactions could be satisfying without being positive, as long as they were not negative. Hence we settled on the term *non-negative* to describe the sort of actions implicated in the need to belong.

Another challenge raised by the reviews was to make the need-to-belong theory look more like a standard motivational theory, insofar as we were claiming it was a standard motivation. This too presented some difficulties, for we did not have a clear sense of what a standard motivational theory should look like. Indeed, our sense was that motivational metatheory was in shambles (as it still is). Social psychologists have good and clear ideas about what cognitive theories should involve, but there is much less of an accepted paradigm for motivation theories. Still, the point was fair, and we thought we ought to try.

Two hallmarks of motivation, substitution and satiation, seemed sufficiently standard that we thought we ought to see whether they could fit into the need to belong. Substitution was easier to address than satiation. A hungry person needs food, and

although the hunger [p. 126 ↓] might be stimulated by a picture of a hamburger, the hunger can be satisfied by salami sandwiches or spaghetti instead. Applying this logic to the need to belong, we reasoned that people must desire at least a certain minimum amount of social connection, and to some extent these bonds could substitute for each other. To be sure, they would most like to keep the bonds they have, but if they lose one relationship, having others could cushion the blow and satisfy some of the same need.

Probably romantic relationships offered the most hostile ground for testing the hypothesis of substitution, because people feel (and are encouraged by cultural messages to feel) that each love relationship is unique and irreplaceable. Yet the data on divorce pointed repeatedly in favor of replaceability. When people divorce, their proneness to all manner of dysfunctional behavior and problems goes up, and when they remarry, it goes back down, indicating that at least in terms of the behavioral and health consequences, the second marriage is an effective substitute for the first. Accounts of romantic heartbreak often feature the tragic theme that the lost love cannot be replaced and no one else will ever do, but in general the person finds someone else who is just as good, if not better (Baumeister and Wotman, 1992; Baumeister et al., 1993). Prisoners, who are forcibly removed from their main relationships, typically form new ones within the prison, sometimes amounting to fully elaborated substitute families, which are important and even cherished at the time but are generally abandoned upon release from prison (Burkhart, 1973; Giallombardo, 1966; Toch, 1977).

The idea of motivational satiation was more difficult to find evidence for (either for or against) than substitution. Satiation means that even though you desire something, at some point you get enough of it and do not want or need any more, at least for a while. Studies of university students were useful in this connection, because students are constantly exposed to other people and could in principle interact with all new people every day, or could form any particular number of stable relationships. They seem mainly to form about four to six main friendships, so that most of their interactions stay within that circle (Wheeler and Nezlek, 1977). Satiation was also suggested in the familiar pattern that when people become deeply involved in an intense relationship, especially a romantic one, they cut back on socializing with friends, as if the one relationship is sufficiently satisfying that they feel less need for the others (e.g., Milardo et al., 1983). In general, we felt we could make a tentative case for satiation, but it

was far from solid. The satiation hypothesis of belongingness theory remains relatively untested even today.

Challenges from Existing Theories

As I said, the assumption of a need to belong was not revolutionary. Our contribution was hardly a discovery of a new drive but rather an enhanced recognition of its reach and power. Several important theories had highly relevant positions, however, and we had to deal with these, as a way of placing our work in the appropriate scholarly context.

One of the best-known motivational theories of the twentieth century was Maslow's hierarchy of needs (e.g., Maslow, 1968). He listed five needs, of which the need to belong was half of one. (He combined it with a need for love.) Moreover, it was the middle one. The middle position meant that belongingness needs had to wait until more basic physiological and safety needs were met, though they would take precedence over later ones such as self-actualization and esteem. Putting it in the middle, thought intuitively reasonable, seemed at odds with our argument that it was basic and powerful. Fortunately, most empirical efforts to test Maslow's theory have concluded that his hierarchy is not correct (e.g., Wahba and Bridwell, 1976). People do pursue satisfactions out of the sequence he specified. [p. 127 ↓] This made it easy to accommodate many findings in which people seemingly jeopardize safety and forego basic animal satisfactions in order to cultivate relationships.

The competing ideas of terror management theory have already been noted. Further work from that group of theorists asserted that avoidance of death was the master motive and that all other motivations are either derivative or secondary (Psychczynski et al., 1997). This challenged our view of the importance and primacy of the need to belong. In their view, the need to belong was at best derivative of the drive to avoid the terrifying thought of death.

We thought there are multiple reasons to give primacy to the need to belong. First, children fear and protest being left alone within the first days of life, but they do not know they will die until quite a few years later. A further and devastating set of findings was reviewed by Leary et al. (1994). They showed that a great many self-presentational

behaviors cause risks or harm to health. The implication is that people do things to make a good impression on other people even if these increase one's risk of death. People sunbathe (even after they have already had skin cancer), ride motorcycles without helmets, smoke cigarettes, and do many other things that fit this category. Clearly, the need to belong sometimes empirically takes precedence over fear of death. The need to belong is not derived from fear of death.

We encountered a challenge of a different sort posed by attachment theory, as originally developed by John Bowlby (e.g., 1969, 1973). His views on the basic drive to connect were quite congenial to our ideas about the need to belong. His approach was derived in important ways from Freud's work, and his understanding of anxiety downplayed the Freudian castration scenario in favor of the more plausible ideas about separation anxiety. Did we have anything new to say beyond Bowlby's attachment theories?

There was a fundamental difference, however. True to his Freudian roots, Bowlby saw the child's relationship to the mother as the root cause of much of what came after, including adult relationships. Even adult employment, in his view, involved seeking and building a relationship to one's supervisor to reenact the erstwhile bond to the mother. In contrast, we saw the child's relationship to the mother as simply one (albeit an early and important) manifestation of the need to belong. In our view, then, the adult's need to belong, and the adult's relationships, were continued signs of the same underlying drive, rather than being a consequence of the connection to the mother.

Extending the Theory: Self-Esteem as Sociometer

As we were finalizing the manuscript and dealing with the reviews, we played around with adding a few extensions and implications to broaden the reach of the idea. At one point I slipped in a sentence or two to suggest that there was possibly a link to self-esteem. When Mark sent me back the draft, he had added a bit more, plus some references to a chapter he was going to publish. I did not realize his thinking was far more advanced than mine on this point. He had begun pulling together the basic ideas and some data for what was to become the sociometer theory. But both of us realized

that the self-esteem issue was worth more than a short paragraph and could turn into a more substantial treatment – one that could take up possibly too much space in this already long paper. We agreed that we would later write another paper developing the implications for self-esteem.

The follow-up paper, which was mainly Leary's work with a few of my thoughts thrown in, addressed the question of why people care about self-esteem (Leary and Baumeister, 2000). Abundant evidence from decades of research indicated that many thoughts, emotions, and behaviors reflected a basic concern with upholding a favorable image of self. In the 1980s this had made [p. 128 ↓] perfect intuitive sense, because people had assumed that high self-esteem produced all manner of beneficial effects. However, the data had gradually failed to support many of those positive assumptions, leaving more and more researchers to wonder why people were so concerned with something (self-esteem) that seemed to have so few pragmatic payoffs (for later review, see Baumeister et al., 2003). One answer was that emotions were tied into self-esteem, and so one might protect self-esteem to prevent unpleasant emotions, but that did not really answer the question: Why were emotional reactions so strongly tied to something that was apparently so useless?

Belongingness offered a powerful context for making sense of self-esteem. Whereas the benefits of self-esteem were dubious and subtle, the benefits of belongingness were clear and strong, as we had found, extending even to life and death. Thus, if self-esteem were tied to belongingness, that would explain why people were so concerned with it.

The link had been established in early studies by Leary et al. (1995). Drops in self-esteem tended to be associated with reductions in belongingness, such as social rejection and the breaking of bonds. Meanwhile, increases in self-esteem were linked to increases in belongingness, such as being accepted for a new relationship or job.

The core idea, then, was that self-esteem served as a sort of mental meter to keep track of belongingness. Humans evolved to be social beings, and so they needed to be connected to each other. Social isolation carried risks to health and wellbeing (plus, obviously, to one's likely success at reproduction), and so the human psyche would benefit from having some sort of alarm go off if the person began to be socially

isolated. Conversely, social acceptance would improve one's prospects for survival and reproduction, and so an inner meter that would register social acceptance in a positive manner would be adaptive. Self-esteem could serve these functions by rising and dropping in response to changes in social belongingness.

Editorial feedback on the Leary and Baumeister (2000) sociometer paper pushed us to sharpen the focus. After all, self-esteem was at best imperfectly correlated with number and closeness of interpersonal attachments; should the correlation not approach perfect unity (well, minus error variance), if that were what self-esteem were all about? Moreover, it seemed clear that emotions already reacted strongly to acceptance and rejection, with bad emotions (anxiety, grief, sadness, jealousy) associated with loss of social connection and plenty of positive ones linked to acceptance, so what was the added value of having self-esteem as a sociometer?

Our solution was to link self-esteem to the anticipated long-term probability of belongingness, not one's current status or momentary changes. Emotion was sufficient to react to momentary changes in belongingness. Self-esteem was rather a relatively stable evaluation, not of how many relationships you have, but how eligible you are to have multiple long-term relationships and other social bonds.

One persuasive indication was that the items on self-esteem scales generally measure four things, all of which seemed quite basic to social acceptance versus rejection. First, all self-esteem scales ask whether you are likable and can get along well with others, which is obviously vital to acceptance. Second, they have items to assess perceived competence, which is important for being employed or otherwise being needed by groups that must perform tasks. Third, many (though not all) scales measure (self-rated) physical attractiveness, which again is a strong determinant of whether others like to be with you. Last, some contain some items assessing moral character, which can be translated into being the sort of person who will conform to the rules of the group such as honesty, fairness, trustworthiness, reliability, and reciprocity. Thus, groups and relationships prefer likable, competent, attractive, and morally well-behaved individuals. Self-esteem is based on those same four criteria.

[p. 129 ↓]

Viewing self-esteem as linked to one's eligibility for, or likelihood of having, good social bonds, rather than counting existing bonds, allowed theoretical room for self-deception. It may be hard for the person with no friends to believe he or she has plenty of friends, but it may be relatively easier to convince oneself that one will soon probably have friends. (The current lack of friends could be dismissed as due to bad luck, special circumstances, or indeed the poor taste of the people around here.) In self-deception, people may concentrate on fooling the meter, as it were: they convince themselves they have traits that will appeal to others without actually having to prove themselves by making these friends.

Another implication is that the focus on eligibility and probability could explain the impostor phenomenon, in which a person might hold low self-esteem despite having multiple friends and other relationships. Such a person might privately fear that, "As soon as they learn what I am really like, they will all desert and abandon me!"

The elaboration of belongingness theory into the sociometer model has implications beyond offering an explanation of self-esteem. Self-esteem is one of the most widely studied constructs in personality, motivation, and self-concept (see Baumeister et al., 2003). To suggest that its primary function involves metering interpersonal relationships and interactions is thus to propose that one of the main intrapsychic structures is driven by interpersonal events.

To me, then, the broadest implication of sociometer theory is this: *inner processes serve interpersonal functions*. In the context of social psychology of recent decades, this is a radical formulation. The prevailing assumption and approach, I believe, treat interpersonal events as a product of intrapsychic ones. That is why so much attention is given to cognition, emotion, and even brain processes: researchers assume that these will offer explanations of what happens between people. I am suggesting that the reverse is more important. What happens between people is of primary importance. What goes on inside them is essentially a set of adaptations to and consequences of the interpersonal events.

The view that inner processes serve interpersonal functions was far from my initial assumptions and theories. But I have been persuaded that it is mostly correct and in powerful, underappreciated ways. If the need to belong is indeed one of the few most

basic, powerful, and pervasive motivations, then inner processes and structures may well have evolved and developed to serve it.

Extending the Theory: Gender Differences in Belongingness

Not long after we published the first paper on the need to belong, I was asked to review a manuscript by Cross and Madson (1997) applying self-construal theory to gender differences. An important cultural difference had been identified by Markus and Kitayama (1991) pertaining to self-construals, which was another term for self-knowledge or self-concepts. As they had stated, Asians tended to think of themselves as interdependent, which meant that their beliefs about themselves focused on how they were related to others, whereas Westerners had more independent self-construals, which meant emphasizing how the individual was different from and stood out from others. Cross and Madson (1997) had proposed that the same distinction could be applied to the differences between men and women in Western culture, namely that men were relatively independent whereas women were more interdependent.

The paper was well documented and well argued, but as I read it I recognized that the thrust of the arguments posed a basic challenge to our theory of the need to belong. In practice, Cross and Madson were suggesting that men did not have the need to belong, or at least not as much as women. Their work was consistent with the newly emerging cultural stereotypes depicting women as [p. 130 ↓] more social than men. For example, Gottman (1994) and others have asserted that women are the experts at relationships and that if men want to have good relationships, they should heed the advice and guidance of their female partners, because men do not generally have the skills and knowledge to make relationships succeed as well as women.

Cross and Madson argued, for example, that the difference in self-construals (and the underlying motivation to be social) was evident in gender differences in aggression. In a nutshell, their argument was that women avoid aggression because it could damage a relationship. If you hit or hurt someone, that person might not want to continue a relationship with you. In contrast, they suggested, men have less desire

and aptitude for relationships, and so they are willing to be aggressive, undeterred by the possible damage to relationships. The theme that men lack the need to belong (especially combined with the impressive amount of evidence that Cross and Madson had compiled) was thus a fascinating challenge to the generality of the theory of the need to belong.

In general, when I review a paper that challenges my work, I try to give it the benefit of the doubt, to compensate for any negative bias that might come from feeling myself attacked. I thought Cross and Madson had made their case reasonably well even though I disagreed with the conclusion. Hence I recommended publishing the paper. I contacted the editor separately, however, and suggested that from my perspective their findings could be interpreted in a completely different manner. The editor, Nancy Eisenberg, encouraged me to write a comment on their manuscript. Kristin Sommer and I started rereviewing the papers Cross and Madson had cited, as well as searching for further evidence.

The theory we developed to account for the gender differences proposed that there are two different spheres of belongingness, reflecting two different ways of being social (Baumeister and Sommer, 1997). Most psychologists and most laypersons give priority to close, intimate relationships, and perhaps women were indeed the experts at these. However, it is also possible to be social in larger groups and networks, and there, we thought, men might seem to be more social. Thus, men could be viewed as social beings with a need to belong, indeed as thoroughly and fundamentally as women, just in a different way.

Return, for a moment, to the issue of gender differences in aggression. There was no disputing the basic finding, replicated over and over, that men are generally more aggressive than women. But does this truly reflect a male indifference to intimate relationship? Abundant evidence was emerging that within close relationships, women were just as aggressive as men (for subsequent review article, see Archer, 2000). Thus, women are slightly more likely than men to physically attack a relationship partner or spouse – everything from a slap in the face to assault with a deadly weapon. Women also perpetrate more child abuse than men, though this was difficult to correct statistically for the vastly greater amount of time they spend with children. Still, the view that women refrain from violence for fear of jeopardizing an intimate bond was

untenable. The preponderance of male violence was not found in family and love relationships but rather between relative strangers and distant acquaintances. Within large networks, men are much more likely than women to come to blows.

We argued, therefore, that the gender difference in aggression reflected not indifference but concern with social relationships. Women care about close relationships and so they are aggressive there when they feel it necessary, the same as men. But women do not care as much about the relatively distant or casual bonds in large groups and networks, whereas men do, and so mainly men are aggressive in those.

Our motivational argument received a huge boost from converging evidence with helping. Cross and Madson (1997) had recognized the extensive data indicating that men help more than women. This was a challenge to their view of women as the more [p. 131 ↓] social beings, and they struggled to reconcile it with their theory. Eventually, they felt it necessary to invoke the feminist clichés that women do not help because they are not socialized to help. But in our response, Sommer and I noted that the evidence of male helpfulness comes mainly from studies with strangers and large groups. Within the family or close relationship, women were plenty helpful, if anything more helpful than men. Thus, the pattern with helping (a major prosocial behavior) was the same as with aggression (a major antisocial behavior): Both men and women do it in close relationships, because both care about these, but mainly men do it in the larger social networks, because men care more about these.

An amusing moment came as Sommer and I were finalizing our manuscript and attended a conference where Leary was also speaking. I gave him the draft of our manuscript and asked him for any thoughts. The next day he told me he had read the first page of the paper, proposing that there are two separate spheres of belongingness, and immediately had the reaction that this was completely wrong ("Oh no!" were his words) and that I was taking our theory in an unfortunate, misguided, and possibly stupid direction. And then, he added, he had read the rest of our paper and become convinced by the mass of evidence that we were right!

The two papers were published together, and the contrasting interpretations stimulated a fair amount of subsequent research. In this case, the evidence has gone in favor of the need-to-belong theory: The social motivations of men extend to large groups

and large networks of shallow relationships much more than women, who emphasize more the intimacy of close relationships (e.g., Benenson and Heath, 2006; Gabriel and Gardner, 1999).

The difference in ways of being social, indeed the different requirements of the two spheres of belongingness, can offer a powerful basis for explaining many of the gender differences in personality (see Baumeister, *in press*). For example, it is widely accepted that women are more emotionally expressive than men. Sharing feelings readily is helpful in close relationships characterized by mutual concern, caring, and support, because the better two people understand each other's feelings, the better they can care for each other. But in a large group, one may have rivals and enemies, and showing all one's feelings can be risky. Likewise, large networks are characterized more than small ones by economic exchange relationships, and a buyer or seller who shows feelings too readily may not be able to get as good a deal as the one who keeps feelings strategically hidden.

Thus, the male pattern of emotional reserve is suited for the large group, whereas the female pattern of emotional expressiveness is more suited to the small group. Similar arguments could be made about many other gender differences, such as agency versus communion, male egotism, emphasis on competence, competitiveness, hierarchy versus egalitarian sentiments, equity versus equality as preferred form of fairness, and rule-based morality versus the so-called "ethic of care" (Baumeister, *in press*).

Not being primarily interested in gender, I had no idea at the time that these ideas would come back later to lead to substantial further reformulations. But these would depend on a very different line of ideas descended from belongingness theory.

Testing the Theory: Effects of Rejection

While finishing up our work on the original review article on the need to belong, I began to wonder about what one might do for laboratory research. It occurred to me that if belongingness were a basic motivation and need, then thwarting it ought to produce a variety of negative, defensive, and/or adaptive reactions. I thought that one might be able to produce strong reactions in the laboratory by manipulating belongingness.

We tried priming the idea of rejection (Sommer and Baumeister, 2002), but this [p. 132 ↓] never worked very well, partly because it was not clear whether the priming made people worry about being rejected or made them feel like rejecting someone else. I wanted to try a live interaction study in which people would choose partners and one person might be told that no one from the group had chosen him or her, which seemed like it ought to be enough to evoke unpleasant emotion and maybe some behavioral reactions. But assembling groups and pulling this off seemed like it would be a hassle, and for that or possibly other reasons, I was unable to persuade any of my graduate students or postdoctorates to try it. Hence the idea remained on the back shelf until Jean Twenge came to work with me. Her background was in personality and she did not have strong commitments as to what to do in social psychology, so she was open to trying the rejection manipulations.

We started with a simple hypothesis about aggression. Rejected people would feel upset over being rejected, and this negative affect would translate into higher aggression. Berkowitz (1989) had recently revisited the frustration-aggression theory, one of the pillars of theory in aggression if not social psychology generally, and he concluded that all negative affect (not just frustration) could drive aggression. It was unclear precisely which negative emotions would arise from being rejected in the laboratory – frustration, disappointment, anxiety, sadness, anger, jealousy – but seemingly it should not matter. Aggression should increase regardless of which emotion mediated it.

The findings from this first investigation on rejection and aggression (Twenge et al., 2001) changed the course of my research program in ways that still affect me to this day. Behavioral manipulations and aggression measures often struggle with uncertain impact, small effects hidden by large variance, and other problems. In contrast, we got very large and reliable effects on behavior. The rejected participants were much more aggressive than the various control groups, even toward people other than those who had rejected them. A parallel study with prosocial behavior yielded equally large effects, though that paper ended up languishing unpublished for the better part of a decade (Twenge et al., 2007) while we ran follow-up studies searching for mediators to satisfy the reviewers. Still, success attracts interest, and soon most of my graduate students wanted to start conducting studies on rejection and social exclusion. New procedures

and measures were developed, and the theory about the effects of the rejection began to fill in.

There were two surprising aspects to these findings. One was the utter lack of emotion. We had reasoned that thwarting a powerful motivation such as the need to belong would elicit emotional distress, which in turn would drive the behavioral effects. Over and over, we found large, reliable behavioral effects, but the emotion was absent. We tried all manner of emotion measures. Participants refused to say that they were upset about our rejection measures. If we did get a significant difference between conditions, which was rare, it was because the socially accepted ones were slightly happy and positive, whereas the rejected ones were neutral. And even these differences never mediated the behavior.

The other surprise was why rejected people should be aggressive, especially if they are not carried away by bad feelings. Presumably rejected people should want to find new connections with others. Being aggressive is counterproductive, because people dislike aggressive persons. The optimal, rational, adaptive response to being rejected should seem to be trying to become nicer. Why do they turn aggressive instead? This question too pushed us to do further work.

New Theorizing about Emotion

Thus, we began our laboratory studies on the need to belong by assuming that emotion would be a crucial, powerful mediator of the effects of rejection, so that emotion would be [p. 133 ↓] the direct cause of behavior. These hypotheses consistently failed to gain support. A recent meta-analysis of nearly 200 studies concluded that rejection conditions generally produce a significant though quite small shift in emotion, away from positive and toward negative – but it typically ends in a neutral state, rather than a negative one (Blackhart et al., *in press*). In essence, rejected persons in the lab say they feel neither good nor bad. And, again, even those small changes in emotion fail to drive behavior. The fallout from that surprising failure has led to efforts to rethink what emotion is for and how it works.

We wondered whether our failure to find that emotion was the mediator and direct cause of behavior was specific to rejection. Rather painstakingly, we waded through all volumes of the field's premier journal, the *Journal of Personality and Social Psychology*, from 1986 to the present (DeWall et al., 2008). The results were somewhat astonishing. Of 201 studies reporting analyses for mediation of behavior by emotion, only 17 percent reported significant results (at the $p = 0.05$ level). Likewise, of 146 studies analyzing for mediation of judgment by emotion, only 18 percent achieved significance. The field's premier journal is thus relatively full of failed analyses for mediation by emotion. Apparently, somebody – authors, reviewers, or editors, or all three – expects emotion to be the direct and immediate cause of behavior. At least, researchers must disprove the default hypothesis of emotional causation before they can draw any other conclusion.

The idea that emotion is the direct cause of behavior is, we think, deeply rooted in psychological theorizing. As we began to read the literature on emotion, we found repeated evocations of the intuitively compelling theory that fear makes one flee, thereby promoting survival, and so the direct causation of behavior by emotion must have served adaptive purposes.

Yet as one takes a skeptical look at the idea that emotion directly causes behavior, it becomes untenable, both in terms of inner gaps and inadequacies and in terms of research findings. Relatively few studies show direct effects of emotion on behavior (see, for example, Schwarz and Clore, 1996, 2007, for reviews). Even when emotion does seem to cause behavior, the addition of proper control groups seems to eliminate those effects (Baumeister et al., 2007).

The eventual result was a major reconceptualization of how emotion is linked to behavior (see Baumeister et al., 2007). Instead of direct causation, emotion seems to operate as a feedback system. After the fact, conscious emotion stimulates reflection, consideration, counterfactual replay, and other cognitive processing that can promote learning for the future. In that respect, emotion can contribute to learning and to eventual behavior change, but it is not well suited for directing behavior in the here and now. Indeed, when people do act based on current emotional state, the results are often self-defeating and maladaptive (e.g., Baumeister and Scher, 1988; Leith and Baumeister, 1996).

Even so, the lack of overt emotion in response to rejection manipulations in our studies remained puzzling. A feedback theory as well as intuition would predict that people would have strong emotional reactions to being rejected, even if these did not serve the purpose of driving behavior. Some reviewers began to grumble that we must be doing something wrong if we failed to elicit emotional distress with our manipulations of social exclusion.

A possible answer came based on several other lines of work. Panksepp (1998, 2005) had proposed dating back to the 1970s that when animals began to evolve in the direction of being social, new inner mechanisms were needed to react to social events. Instead of creating wholly new ones, evolution piggybacked the social functions onto the existing structures that registered pleasure and pain. This raised the possibility that social trauma might produce effects resembling physical trauma, including a brief numbness in a state of shock that would allow the animal to [p. 134 ↓] finish the battle or complete its escape despite being injured, without being crippled by intense pain. Might social rejection produce the same sort of numbness?

A remarkable article by MacDonald and Leary (2005) reviewed evidence, mostly from the animal literature, on the effects of social exclusion and rejection on physical pain. Sure enough, animals appeared to go numb to physical pain as a result of being excluded by the social group. For example, rat pups ejected from the litter become relatively insensitive to pain.

Intrigued by this possibility, we purchased a device for administering and calibrating physical pain and began to measure how reactions changed when we thwarted the need to belong. To our surprise, pain sensitivity diminished quickly and substantially in response to our laboratory manipulations of rejection. Crucially, and consistent with Panksepp's theory, we found that emotional responses diminished in step with the analgesia.

Apparently, social exclusion causes the body to go numb, both physically and emotionally. The emotional numbness extended even to things bearing no meaningful connection to the rejection, such as predicting one's emotional state as a function of the outcome of next month's big football game, or empathizing with the plight of a fellow student whose leg was broken (DeWall and Baumeister, 2006). Thus, rejected people

did not even show the widely replicated patterns of affective forecasting (i.e., predictions of one's own future emotional reactions), in which people overestimate their emotional responses (Wilson and Gilbert, 2003).

Still, it seemed intuitively likely that people would be somewhat upset by being rejected in our studies. Perhaps the shock and numbness response suppressed conscious emotion, but there still ought to be nonconscious reactions. We therefore began to test the effects of exclusion manipulations on implicit measures of emotion. One idea, for example, was that people would react to exclusion by actively suppressing their emotions, so that they would report no emotion, but their implicit reactions would exhibit considerable distress.

The results of this line of investigation shocked us so much that we ended up replicating them over and over, using different measures and different manipulations (but always getting the same result). Social exclusion did not increase nonconscious distress. It did, however, increase nonconscious positive emotions. Participants who underwent our rejection manipulations showed an increase in the accessibility of affectively positive thoughts and associations. Negative thoughts and associations did not differ between rejection and control conditions (Twenge et al., 2007).

Apparently, when bad things happen, the human psyche copes in the short run by some biochemical response (presumably involving the release of opioids) to create numbness so that the immediate emotional reaction is dampened and forestalled. Meanwhile, the nonconscious mind begins searching for happy thoughts and associations. Thus, when the numbness wears off and the emotion can begin to be felt, it is somewhat reduced by the positive thoughts that have been dragooned in the mean time.

Human Nature and Culture

The theory about the need to belong led slowly toward a rethinking of human nature itself. The impetus for this was a book that sought to package psychology's message for a broader, interdisciplinary audience. It was dismaying to realize that when scholars in other fields want to use psychology in their work, they still sometimes rely on Freud's

system. I thought it would be useful to read through our field's laboratory findings and compile a new, integrated, systematic account of human nature.

The project proceeded piecemeal. I had no overarching ideas of how the human psyche was put together. I organized my search [p. 135 ↓] based on questions, not answers. The questions were the ones that I thought scholars in other fields would want to know from psychology: What do people want? How do they think? What do their emotions do? How do people act? How do they interact? And how do they grow up and turn out? I amassed plenty of specific facts and conclusions but the broad, integrative insight I sought eluded me until very late in the project.

A long tradition of psychological and even philosophical theory (through Freud, 1930, and back to the philosopher Rousseau) had emphasized the basic conflicts between individual and society. I had accepted and written about this conflict (Baumeister, 1986a, 1987). Yet the idea of basic hostility between person and society did not square well with the outlines of the human psyche I had documented. Instead, the patterns of wanting, thinking, feeling, and acting seemed designed to facilitate interactions with culture, rather than avoid it or cope with its deficiencies.

This brought me back to the lesson from sociometer theory: *inner processes serve interpersonal functions*. Perhaps the main parts of the human psyche, especially its uniquely human parts, were chiefly designed to enable people to interact with each other. The basic fact of human existence is not conflict but rather cooperation between individual and society.

The statement that humans are social animals had been made by thinkers from Aristotle to Aronson, and the phrase "social animal" has become standard in social psychology. It is true that humans are social animals, but I felt this did not go far enough. There are many social animals. What sets humans apart is the highly advanced form of social life we have developed, specifically culture. And so my organizing principle became the idea that humans are cultural animals (Baumeister, 2005).

Was it plausible? A long tradition has assumed that people evolved physically first and then invented culture. I was suggesting, instead, that the progressively advancing demands of ever more complex social life may have helped pull physical evolution to

develop greater capabilities, indeed the features that make us human. I discovered that other theorists had begun to argue for coevolution of culture and physical capabilities (Boyd and Richerson, 1985).

Meanwhile, ethologists were beginning to argue that rudimentary forms of culture could be observed in multiple other species (De Waal, 2001). If culture existed in other species, then it may have been on the planet before humans did. This brought a crucial implication that no one seemed to have noted. If culture were here before humankind, then culture might easily have been part of the selection environment in which humans evolved. That meant that culture could well have shaped the evolution of human capabilities. Culture could have shaped nature.

A remarkable program of research by Dunbar (1993, 1998) had produced revolutionary ideas about human intelligence. Intelligence has a special place in theories of human nature because humans have named our own species (*Homo sapiens*) after our intelligence. Like most scholars, I had assumed that human intelligence had evolved for solving problems in the physical environment, such as finding food and outsmarting predators. Dunbar tested a variety of such theories and found no support for them. Instead, he found that brain size was linked to the size and complexity of social networks, which he dubbed the “social brain” hypothesis. Although he had not included humans in his studies, the implications are there in spades. Our intelligence was not developed for outsmarting the bears and rabbits, but for understanding each other.

All this cast the need to belong in a new light. To be sure, plenty of animals have a sort of herd instinct that prompts them to be together. They are safer that way. But for humans, social interaction goes beyond a herd instinct. Social interaction (including culture) is humankind's biological strategy. [p. 136 ↓] Humans form relationships, not just intimate bonds but larger social structures with diverse roles, shared information, and even market economies, because these are what have brought the remarkable biological success of our species.

Culture is thus a new way of being social. The human need to belong entails more than just wanting to be near other people. Combined with the improved information processing capacity, it enables people to make culture. And of course the information

processing capacity itself is mainly used for dealing with social information also, especially including work and employment.

The cultural animal perspective does put me once again out of step with the prevailing trends in our field, and so I seem condemned to remain a contrarian. These days, the biggest new thrust of research is to explore the brain. I support brain research, but I also think that most psychologists greatly overestimate what is to be learned from it. My impression is that many psychologists believe that studying the brain will reveal the fundamental root causes and ultimate explanations of human behavior. That is why they are willing to suspend research on observing actual behavior in order to focus on brain imaging. They think social behaviors and interactions follow from what happens in the brain.

I think it is the other way around. The brain will gradually be revealed to be a mere switchboard, not the ultimate cause. The ultimate causes lie in social interaction, and the brain evolved as it did in order to facilitate those interactions. The brain is the hardware for inner processes, and inner processes serve interpersonal functions. My prediction is that within a couple decades, after we have mapped out all the various sites and activities in the brain, we will realize that we have not explained social behavior – that in fact we need to take another earnest look at social interaction in order to understand why the brain is the way it is. Relationships are not created by the brain; rather, the brain was created to serve relationships.

Men, Women, and Culture

Understanding the centrality of culture to human nature pointed the way toward a reconsideration of gender differences. Ultimately, these lines of thought could offer the basis for a sweeping reconceptualization of gender politics, which in recent decades has become dominated by the view that men oppressed women as a main theme of world history.

Earlier I sketched out the theory and evidence about two spheres of belongingness. The core idea was that male and female psyches had slightly different versions of the need

to belong, with the women focused primarily on close, intimate relationships and the men orienting more toward larger networks of shallower relationships.

Cultural progress depended on large networks of shallow relationships, because these offer fertile opportunities for sharing information, disseminating innovations, and creating wealth via economic trade (see Baumeister, 2005). Let us assume that prehistoric humans divided their social worlds to some extent into male and female spheres (possibly based on the hunter-gatherer and other task differentiations) and that the separate social worlds took the form favored by most people in them. The women's sphere would be one of close, intimate relationships and pair bonds, whereas the male sphere would develop larger networks of shallower relationships. The male sphere would thus offer a more viable ground than the female sphere for culture to develop.

Hence wealth, knowledge, and power were largely created in the male sphere. This is what accounted for the emergence of gender inequality. It was neither that men had superior talents and abilities, nor that men banded together to conspire to oppress women. (These have been the two main explanations.) It was simply that the male sphere made progress while the women's sphere remained essentially static, because of the different types of social relationships favored by the respective genders. The men's large networks [p. 137 ↓] of shallow relationships are not as satisfying or as nurturant as the intimate bonds cultivated by women, but they are conducive to the requirements of cultural progress, such as dissemination of information, competitive innovation, and economic exchange.

I see this line of thought as representing the opportunity for a truce in the so-called battle of the sexes (see Baumeister, 2007, *in press*). It is not necessary to regard one gender as superior to other, nor to postulate that men and women were fundamentally enemies and rivals, with men becoming oppressors and women victims. Rather, by inclination and social structure, men and women made separate contributions to the group's and indeed the species' welfare. Women provided the vital relationships that enabled the species to reproduce itself and thus to survive, whereas men competed in larger networks and gradually produced most of the advances that made cultural progress. Men and women have been partners more than enemies. Culture advanced not by men banding together against women, but mainly by groups of men competing against other groups of men.

Thus, based on the theory of the need to belong, one can formulate a kinder and gentler account of gender politics. Instead of seeing men as oppressors and women as victims, one can understand the genders as working together. Men's relationships may be less emotionally nurturant and intimately satisfying as female relationships, but they provide alternative advantages. Perhaps the most important of these advantages is that they foster the kinds of interactions and exchanges that make culture possible and enable it to make progress.

Applications and Implications

Although my goal in this and other theories has been to understand people rather than to change the world, there are assorted implications and potential applications. Most obviously, the findings about the bad effects of rejection indicate the importance of fostering an inclusive society and enabling people to find ways to feel connected with others. Socially excluded people have responded in our laboratory studies with increased aggression, reduced helpfulness, and cooperation with poor self-control, with self-defeating behavior, and with impairments in intellectual performance. Indeed, sociologists have frequently observed that excluded minority groups in many societies exhibit just such socially maladaptive patterns of poor school and intellectual performance, violence and crime, and lack of prosocial activity. Although these behaviors are sometimes regarded as the basis for excluding them, exclusion appears to be a causal factor in exacerbating and even producing those patterns.

Hence one practical implication for societies and organizations is to ensure that all, including those at the bottom, are encouraged to feel accepted and included. It may be tempting for those at the top of the hierarchy to feel that they matter the most, but the destructive responses by those who feel excluded can cause significant costs to all.

Another implication for clinicians and others is that people seem to need at least a minimum number of (perhaps four) close relationships before the need is reasonably satisfied. Simply counting the number of people a person feels close to may be a useful first step in understanding that person's life, world, and prognosis. As to what people are working on in life, those with too few relationships will likely give high priority to finding new bonds, whereas those who are well connected may scarcely give that project

another thought. In this modern era of small families, single-person households, and transient relationships, more and more time and energy are likely to be devoted to the task of making social connections.

The new understanding of gender politics also suggests important changes for society. The societal quest to eliminate all gender differences in achievement is likely to remain futile. Assuming that all gender differences [p. 138 ↓] stem from male oppression and can be cured by benefiting women is likewise going to lead to disappointment. Meanwhile, society's efforts to change everything for the benefit of women and girls may begin to have harmful effects on boys and men, as is already abundantly evident in American schools. Instead of seeing men and women as basically identical sorts of beings who are political enemies, it may be preferable to regard them as slightly different sorts of creatures who make good partners.

Conclusion

The theory of the need to belong has evolved from a simple insight about desiring togetherness into a large complex of interrelated ideas. These address fundamental questions about human nature, culture, gender, emotion, and the relation between inner and interpersonal processes.

The work is far from complete, and further advances seem likely. Some aspects of the basic theory, such as whether people become satiated when they have enough relationships, remain open questions for further work. But given how the theory has developed from relatively unforeseen connections and findings, it seems reckless to predict where it will go next.

Ultimately, it seems vital for psychology to rediscover an appreciation for what happens between people. What happens inside them is important also, but often it happens because of what happens between them. Inner processes serve interpersonal functions. The need to belong is a fundamental and powerful fact about human nature.

References

- Archer, J. Sex differences in aggression between heterosexual partners: A meta-analytic review . Psychological Bulletin , 126, 697–702. (2000) <http://dx.doi.org/10.1037/0033-2909.126.5.697>
- Baumeister, R.F. (1986a) Identity: Cultural Change and the Struggle for Self . New York: Oxford University Press.
- Baumeister, R.F. (ed.) (1986b) Public Self and Private Self . New York: Springer-Verlag. <http://dx.doi.org/10.1007/978-1-4613-9564-5>
- Baumeister, R.F. How the self became a problem: A psychological review of historical research . Journal of Personality and Social Psychology , 52, 163–176. (1987) <http://dx.doi.org/10.1037/0022-3514.52.1.163>
- Baumeister, R.F. (2005) The Cultural Animal: Human Nature, Meaning, and Social Life . New York: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780195167030.001.0001>
- Baumeister, R.F. (2007) Is there anything good about men? Invited address presented to the American Psychological Association, San Francisco, August.
- Baumeister, R.F. (2010) Is There Anything Good About Men? How Cultures Flourish by Exploiting Men . New York: Oxford University Press.
- Baumeister, R.F., Campbell, J.D., Krueger, J.I. and Vohs, K.D. Does high self-esteem cause better performance interpersonal success, happiness, or healthier lifestyles? Psychological Science in the Public Interest , 4, 1–44. (2003) <http://dx.doi.org/10.1037/0033-2909.90.2.245>
- Baumeister, R.F. and Leary, M.R. The need to belong: Desire for interpersonal attachments as a fundamental human motivation . Psychological Bulletin , 117, 497–529. (1995) <http://dx.doi.org/10.1037/0033-2909.117.3.497>

Baumeister, R.F. and Scher, S.J. Self-defeating behavior patterns among normal individuals: Review and analysis of common self-destructive tendencies . Psychological Bulletin , 104, 3–22. (1988) <http://dx.doi.org/10.1037/0033-2909.104.1.3>

Baumeister, R.F. and Sommer, K.L. What do men want? Gender differences and two spheres of belongingness: Comment on Cross and Madson (1997) . Psychological Bulletin , 122, 38–44. (1997) <http://dx.doi.org/10.1037/0033-2909.122.1.38>

Baumeister, R.F. and Tice, D.M. Anxiety and social exclusion . Journal of Social and Clinical Psychology , 9, 165–195. (1990) <http://dx.doi.org/10.1521/jscp.1990.9.2.165>

Baumeister, R.F., Vohs, K.D., DeWall, C.N. and Zhang, L. How emotion shapes behavior: Feedback, anticipation, and reflection, rather than direct causation . Personality and Social Psychology Review , 11, 167–203. (2007) <http://dx.doi.org/10.1177/1088868307301033>

Baumeister, R.F. and Wotman, S.R. (1992) Breaking Hearts: The Two Sides of Unrequited Love . New York: Guilford Press.

Baumeister, R.F., Wotman, S.R. and Stillwell, A.M. Unrequited love: On heartbreak, anger, guilt, scriptlessness, and humiliation . Journal of Personality and Social Psychology , 64, 377–394. (1993) <http://dx.doi.org/10.1037/0022-3514.64.3.377>

Becker, E. (1973) The Denial of Death . New York: Free Press.

Benenson, J.F. and Heath, A. Boys withdraw more in one-on-one interactions, whereas girls withdraw more in groups . Developmental Psychology , 42, 272–282. (2006) <http://dx.doi.org/10.1037/0012-1649.42.2.272>

Berkowitz, L. Frustration-aggression hypothesis: Examination and reformulation . Psychological Bulletin , 106, 59–73. (1989) <http://dx.doi.org/10.1037/0033-2909.106.1.59>

Blackhart, G.C., Knowles, M.L., Nelson, B.C. and Baumeister, R.F. (in press) Rejection elicits emotional reactions but neither causes immediate distress nor lowers self-

esteem: A meta-analytic review of 192 studies on social exclusion . Personality and Social Psychology Review .

Bowlby, J. (1969) Attachment and Loss. Vol. 1: Attachment . New York: Basic Books.

Bowlby, J. (1973) Attachment and Loss. Vol. 2: Separation Anxiety and Anger . New York: Basic Books.

Boyd, R. and Richerson, P.J. (1985) Culture and the Evolutionary Process . Chicago: University of Chicago Press.

Burkhart, K. (1973) Women in Prison . Garden City, NY: Doubleday.

Cross, S.E. and Madson, L. Models of the self: Self-construals and gender . Psychological Bulletin , 122, 5–37. (1997) <http://dx.doi.org/10.1037/0033-2909.122.1.5>

De Waal, F. (2001) The Ape and the Sushi Master: Cultural Reflections of a Primatologist . New York: Basic Books.

DeWall, C.N. and Baumeister, R.F. Alone but feeling no pain: Effects of social exclusion on physical pain tolerance and pain threshold, affective forecasting, and interpersonal empathy . Journal of Personality and Social Psychology , 91, 1–15. (2006) <http://dx.doi.org/10.1037/0022-3514.91.1.1>

DeWall, C.N., Baumeister, R.F. and Bushman, B.J. (2008) How frequently does emotion mediate behavior: A survey of articles in the Journal of Personality and Social Psychology . Manuscript submitted for publication.

Dunbar, R.I.M. Coevolution of neocortical size, group size, and language in humans . Behavioral and Brain Sciences , 16, 681–694. (1993) <http://dx.doi.org/10.1017/S0140525X00032325>

Dunbar, R.I.M. The social brain hypothesis . Evolutionary Anthropology , 6, 178–190. (1998) <http://dx.doi.org/10.1002/%28SICI%291520-6505%281998%296:5%3C178::AID-EVAN5%3E3.0.CO;2-8>

Freud, S. (1930) Civilization and Its Discontents (J. Riviere, ed. , trans.). London: Hogarth Press.

Gabriel, S. and Gardner, W.L. Are there 'his' and 'her' types of interdependence? The implications of gender differences in collective and relational interdependence for affect, behavior, and cognition . *Journal of Personality and Social Psychology* , 75, 642–655. (1999) <http://dx.doi.org/10.1037/0022-3514.77.3.642>

Gerstel, N. and Gross, H. Commuter marriages: A review . *Marriage and Family Review* , 5, 71–93. (1982) http://dx.doi.org/10.1300/J002v05n02_05

Gerstel, N. and Gross, H. (1984) Commuter Marriage: A Study of Work and Family . New York: Guilford Press.

Giallombardo, R. (1966) Society of Women: A Study of a Women's Prison . New York: Wiley.

Gottman, J.M. (1994) What Predicts Divorce? Hillsdale, NJ: Erlbaum.

Greenberg, J. and Pyszczynski, T. Compensatory self-inflation: A response to the threat to self-regard of public failure . *Journal of Personality and Social Psychology* , 49, 273–280. (1985) <http://dx.doi.org/10.1037/0022-3514.49.1.273>

Greenberg, J., Pyszczynski, T. and Solomon, S. (1986) The causes and consequences of self-esteem: A terror management theory . In R. Baumeister (ed.), *Public and Private Self* . New York: Springer-Verlag.

Harrison, A.A. and Connors, M.M. (1984) Groups in exotic environments . In L. Berkowitz (ed.), *Advances in Experimental Social Psychology* , 18, 49–87. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960142-2>

Kunz, P.R. and Woolcott, M. Season's greetings: From my status to yours . *Social Science Research* , 5, 269–278. (1976) <http://dx.doi.org/10.1016/0049-089X%2876%2990003-X>

Latané, B., Eckman, J. and Joy, V. Shared stress and interpersonal attraction . *Journal of Experimental Social Psychology Supplement* , 1, 80–94. (1966)

Leary, M.R. Responses to social exclusion: Social anxiety, jealousy, loneliness, depression, and low self-esteem . *Journal of Social and Clinical Psychology* , 9, 221–229. (1990) <http://dx.doi.org/10.1521/jscp.1990.9.2.221>

Leary, M.R. and Baumeister, R.F. (2000) The nature and function of self-esteem: Sociometer theory . In M. Zanna (ed.), *Advances in Experimental Social Psychology* , 32, 1–62. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2800%2980003-9>

Leary, M.R., Tambor, E.S., Terdal, S.K. and Downs, D.L. Self-esteem as an interpersonal monitor: The sociometer hypothesis . *Journal of Personality and Social Psychology* , 68, 518–530. (1995) <http://dx.doi.org/10.1037/0022-3514.68.3.518>

Leary, M.R., Tchividjian, L.R. and Kraxberger, B.E. Self-presentation can be hazardous to your health: Impression management and health risk . *Health Psychology* , 13, 461–470. (1994) <http://dx.doi.org/10.1037/0278-6133.13.6.461>

Leith, K.P. and Baumeister, R.F. Why do bad moods increase self-defeating behavior? Emotion, risk taking, and self-regulation . *Journal of Personality and Social Psychology* , 71, 1250–1267. (1996) <http://dx.doi.org/10.1037/0022-3514.71.6.1250>

Lynch, J.J. (1979) *The Broken Heart: The Medical Consequences of Loneliness* . New York: Basic Books.

MacDonald, G. and Leary, M.R. Why does social exclusion hurt? The relationship between social and physical pain . *Psychological Bulletin* , 131, 202–223. (2005) <http://dx.doi.org/10.1037/0033-2909.131.2.202>

Markus, H.R. and Kitayama, S. Culture and the self: Implications for cognition, emotion, and motivation . *Psychological Review* , 98, 224–253. (1991) <http://dx.doi.org/10.1037/0033-295X.98.2.224>

Maslow, A.H. (1968) *Toward a Psychology of Being* . New York: Van Nostrand.

McLeod, E. (1982) Women Working: Prostitution Today . London: Croom Helm.

Milardo, R.M., Johnson, M.P. and Huston, T.L. Developing close relationships: Changing patterns of interaction between pair members and social networks . Journal of Personality and Social Psychology , 44, 964–976. (1983) <http://dx.doi.org/10.1037/0022-3514.44.5.964>

Panksepp, J. (1998) Affective Neuroscience: The Foundations of Human and Animal Emotions . New York: Oxford University Press.

Panksepp, J. Why does separation distress hurt? Comment on MacDonald and Leary (2005) . Psychological Bulletin , 131, 24–230. (2005) <http://dx.doi.org/10.1037/0033-2909.131.2.224>

Pyszczynski, T., Greenberg, J. and Solomon, S. Why do we need what we need? A terror management perspective on the roots of human social motivation . Psychological Inquiry , 8, 1–20. (1997) http://dx.doi.org/10.1207/s15327965pli0801_1

Schwarz, N. and Clore, G.L. (1996) Feelings and phenomenal experiences . In E.T. Higgins, ed. and A. Kruglanski (eds), Social Psychology: Handbook of Basic Principles , pp. 433–465. New York: Guilford Press.

Schwarz, N. and Clore, G.L. (2007) Feelings and phenomenal experiences . In E.T. Higgins, ed. and A. Kruglanski (eds), Social Psychology: Handbook of Basic Principles , 2nd Edition. New York: Guilford Press.

Sommer, K.L. and Baumeister, R.F. Self-evaluation, persistence, and performance following implicit rejection: The role of trait self-esteem . Personality and Social Psychology Bulletin , 28, 926–938. (2002)

Symanski, R. (1980) Prostitution in Nevada . In E. Muga (ed.), Studies in Prostitution , pp. 246–279. Nairobi: Kenya Literature Bureau.

Toch, H. (1977) Living in Prison: The Ecology of Survival . New York: Free Press.

Twenge, J.M., Baumeister, R.F., DeWall, C.N., Ciarocco, N.J. and Bartels, J.M. Social exclusion decreases prosocial behavior . *Journal of Personality and Social Psychology* , 92, 56–66. (2007) <http://dx.doi.org/10.1037/0022-3514.92.1.56>

Twenge, J.M., Baumeister, R.F., Tice, D.M. and Stucke, T.S. If you can't join them, beat them: Effects of social exclusion on aggressive behavior . *Journal of Personality and Social Psychology* , 81, 1058–1069. (2001) <http://dx.doi.org/10.1037/0022-3514.81.6.1058>

Wahba, M.A. and Bridwell, L.G. Maslow reconsidered: A review of research on the need hierarchy theory . *Organizational Behavior and Human Performance* , 15, 212–240. (1976) <http://dx.doi.org/10.1016/0030-5073%2876%2990038-6>

Wheeler, L. and Nezlek, J. Sex differences in social participation . *Journal of Personality and Social Psychology* , 35, 742–754. (1977) <http://dx.doi.org/10.1037/0022-3514.35.10.742>

Wilson, T.D. and Gilbert, D.T. (2003) Affective forecasting . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 35, 345–411. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2803%2901006-2>

<http://dx.doi.org/10.4135/9781446249222.n32>

Handbook of Theories of Social Psychology

Sociometer Theory

Contributors: Mark R. Leary

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Sociometer Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n33>

Print pages: 141-160

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n33>

[p. 141 ↓]

Chapter 33: Sociometer Theory

Mark R. Leary

Abstract

Sociometer theory proposes that self-esteem is a psychological gauge of the degree to which people perceive that they are relationally valued and socially accepted by other people. In conceptualizing self-esteem as the output of a system that monitors and responds to interpersonal acceptance and rejection, sociometer theory differs from most other explanations of self-esteem in suggesting that people neither need self-esteem nor are motivated to pursue it for its own sake. Rather, according to the theory, when people do things that appear intended to protect or increase their self-esteem, their goal is usually to protect and enhance their relational value and, thus, increase the likelihood of interpersonal acceptance. This chapter reviews previous conceptualizations of self-esteem, examines the development and details of sociometer theory, and reviews research evidence that supports three central predictions of the theory – that acceptance and rejection influence state self-esteem, state self-esteem relates to perceived social acceptance, and trait self-esteem reflects people's perceptions of their general acceptability or relational value. Common objections to sociometer theory are also addressed, along with implications of the theory for clinical and societal interventions.

Introduction

Self-esteem ranks among the most widely studied constructs in the social and behavioral sciences. More than 30,000 articles and chapters have been published that include the term, "self-esteem," in their title or abstract, and hundreds of thousands more have undoubtedly included some mention of self-esteem. Such a level of scholarly

interest suggests that theorists, researchers, and practitioners have regarded self-esteem as an exceptionally important construct, one that helps to explain a great deal of human thought, emotion, and behavior. Fascination with self-esteem has crept into the public mind as well. Not only is self-esteem a frequent topic of popular books and magazine articles, but also most people seem to believe that self-esteem is an essential ingredient in happiness, wellbeing, and success.

Oddly, despite all of the attention, throughout most of the twentieth century, few theorists made an effort to explain why self-esteem [p. 142 ↓] is psychologically important, what it actually does, or why it is worthy of so much attention. Sociometer theory was offered in an effort to explain the function of self-esteem and to account for its known relationships with an array of psychological phenomena.

Intellectual Background

Early Ideas about Self-Esteem

Within psychology, self-esteem made its first notable appearance in James' (1890) textbook, *Psychology*. In his chapter on the self, James discussed the feelings that are associated with "self-appreciation" (by which he meant self-evaluation). According to James, the primary causes of people's self-feelings involve their "actual success or failure, and the good or bad actual position one holds in the world" (1890: 182). Yet, he also noted that self-esteem depends not only on people's actual outcomes in life but also on how they judge their outcomes relative to their aspirations. He offered a simple formula suggesting that self-esteem reflects the ratio of one's successes to one's pretensions and observed that people can change their own self-feelings because "such a fraction may be increased as well by diminishing the denominator as by increasing the numerator" (1890: 187). Although James suggested that people wish to feel good rather than bad about themselves, he did not discuss the functions that self-esteem might serve.

Most subsequent writers have followed James' lead, offering ideas about the causes and consequences of self-esteem without entertaining the question of why people have

the capacity for self-esteem or what functions it might serve. For example, Rogers (1959) viewed self-esteem as a consequence of the degree to which people receive positive regard from other people. He suggested that people instinctively need both positive regard (e.g., love, affection, attention, nurturance) from other people and positive self-regard (i.e., self-esteem) in order to become fully functioning. However, because positive regard often depends on behaving in a particular way or being a particular kind of person, people bend themselves in ways that are contrary to their natural inclinations and best interests to obtain it. In the process, self-regard also becomes conditional so that people feel good about themselves only when they meet other people's standards rather than when they actualize their own potentials. Although people who act against their own interests, values, personality, and inclinations probably fare more poorly than those who behave congruently with their inclinations, Rogers and other humanistic theorists did not articulate the role of self-esteem in this process, which can be explained in terms of conformity to social pressures that are incongruent with one's inclinations without invoking the concept of self-esteem at all.

Once the assumption that people need self-esteem became established, theorists began to explain certain behaviors in terms of people's quest for self-esteem. Branden (1969) in particular stressed the role that the basic "need" for self-esteem plays in human behavior. More than most theorists, Branden's view of self-esteem was particularly nonsocial, with the basic "pillars" of self-esteem residing in intrapsychic processes involving self-awareness, integrity, personal responsibility, and being honest with oneself (Branden, 1983).

Functional Explanations of Self-Esteem

Based on the work on Rogers (1959), Branden (1969), and others, the notion that people have a need for self-esteem became a well-entrenched assumption, but few writers addressed the broader question of why people need self-esteem, what it does, or why deficiencies in self-esteem seem to be associated with poor mental health. However, during the 1970s and 1980s, three broad functional perspectives on the nature of self-esteem emerged.

[p. 143 ↓]

Self-Esteem Promotes Adaptive Behavior and Psychological Wellbeing

The most widely espoused explanation of self-esteem's function suggested that self-esteem promotes psychological wellbeing and success (Bandura, 1977; Greenwald, 1980; Taylor and Brown, 1988). Evidence that self-esteem correlates with happiness and success led many to conclude that self-esteem facilitates positive outcomes. For example, research showed that people with high self-esteem perform better after failure than people with low self-esteem and are also more likely to persevere in the face of obstacles (Perez, 1973; Shrauger and Sorman, 1977). Furthermore, studies showed that trait self-esteem was associated with greater success in academics (Bowles, 1999; Hansford and Hattie, 1982) and social life (Glendinning and Inglis, 1999; LePine and Van Dyne, 1998). Such findings led many to conclude that high self-esteem promotes a variety of positive outcomes and that people seek self-esteem for these benefits. In general, theorists did not explain why self-esteem is related to such outcomes, but at the time the connection seemed indisputable (Taylor and Brown, 1988).

However, this explanation of self-esteem has at least three weaknesses. First, it does not explain why merely feeling good about oneself leads to positive outcomes. Many of these explanations confused self-esteem (valenced feelings about oneself) with self-efficacy (the belief that one is capable of achieving desired outcomes). Although beliefs regarding one's own efficacy promote motivation, perseverance, and positive emotion (Bandura, 1977; Cervone et al., 2006), whether merely feeling good about oneself enhances goal-directed behavior is less clear.

Second, reviews of the literature on the relationship between self-esteem and positive outcomes – including achievement, interpersonal success, adjustment, and social problems – showed that the relationships are far weaker than typically assumed (Baumeister et al., 2003; Mecca et al., 1989). This fact does not dispute that high self-esteem might be associated with beneficial outcomes, but it raises questions regarding whether the function of the self-esteem motive is to facilitate outcomes such as these.

Third, many theorists fell victim to the fallacy of inferring causality from correlation. Virtually all evidence showing that self-esteem is associated with positive outcomes is correlational, yet many writers concluded that self-esteem was the causal factor. Few writers entertained the possibility that self-esteem was caused by positive outcomes, such as academic or social success or that self-esteem, positive emotions, and behavioral effectiveness are effects of some other process (Baumeister et al., 2003).

Self-Esteem Signals Dominance or Status

A second perspective was proposed in different ways by Tedeschi and Norman (1985) and Barkow (1980) who suggested that self-esteem is an indicator of the degree to which a person has influence or dominance over others. Tedeschi and Norman (1985) suggested that feelings of self-esteem indicate to people that they have influence or power over others and, thus, people seek self-esteem because it is a reinforcer that is associated "with the facilitation of social influence and the attainment of rewards" (1985: 310). Put differently, people seek self-esteem because the behaviors that lead to high self-esteem are associated with having influence over others.

Barkow (1980) rooted a similar explanation in evolutionary ethology. To the extent that being dominant in one's social group would have been associated with greater access to mates and other resources during evolutionary history, the tendency to monitor and increase one's relative social standing would confer fitness benefits. As human beings developed the capacity for self-relevant thought, they may have developed mechanisms for evaluating and enhancing their relative dominance. To the extent that self-esteem is associated with having status and being dominant, people may seek self-esteem not because it has any inherent value but rather because it connotes dominance.

[p. 144 ↓]

Little research has directly examined the hypothesis that self-esteem monitors dominance, although studies show that dominant people tend to have higher trait self-esteem than less dominant people (Heaven, 1986; Leary et al., 2001; Raskin et al., 1991). Further more, two experimental studies that manipulated participants' perceptions of their dominance (by varying the degree to which they believed that

other participants wanted them to be group leader) showed that perceived dominance influences state self-esteem. Although the notion that self-esteem is linked to the exercise of status and dominance may have some merit, these explanations do not fully account for all of the factors that influence self-esteem or explain all instances in which people appear to seek to enhance their self-esteem (Leary et al., 2001).

Self-Esteem Lowers Existential Anxiety

Terror management theory focuses on the psychological and behavioral tactics that people use to ward off the existential terror that they experience at the prospect of their own death and annihilation (Solomon et al., 1991). According to terror management theory, self-esteem lowers paralyzing terror because death-related anxiety is reduced by the belief that one is a valuable individual in a meaningful universe, and high self-esteem is related to the belief that one is living up to important cultural values (Pyszczynski et al., 2004). Thus, people with high self-esteem feel assured that they will achieve immortality – either literally in terms of going to heaven or being reincarnated, or symbolically in that their impact will remain after they die. In either case, self-esteem buffers them against the anxiety they would otherwise feel.

Like dominance theory, terror management theory moved away from the longstanding assumption that people need self-esteem for its own sake to the idea that people seek self-esteem because it serves an important function. Furthermore, like dominance theory, terror management theory provides a cogent explanation for why self-esteem is linked to one's social standing in other people's eyes: the anxiety-buffering mechanism is effective only to the extent that people believe that they are a person-of-value according to the standards of their culture.

As a broad theory of the pervasive effects of existential concerns on human behavior, terror management theory has much to recommend it, and a great deal of research has supported its central propositions (Pyszczynski et al., 2004). However, its explanation of self-esteem may have the weakest empirical support. Although self-esteem is associated with anxiety as the theory predicts, this finding can be explained by virtually every other theory of self-esteem as well (Leary, 2004).

Key Issues

Each of these perspectives offers insights regarding the nature of self-esteem, but each leaves many unanswered questions, and none addresses all known patterns of findings. Two issues remain particularly problematic. First, most theorists have assumed that self-esteem reflects a private self-evaluation in which people compare themselves with either their personal standards or an idealized sense of self. Yet, research shows that self-esteem is as strongly related to people's beliefs about other people's perceptions and evaluations of them as to their own self-perceptions, if not more so (Leary and Baumeister, 2000). If self-esteem reflects people's private self-evaluations, divorced from interpersonal concerns, then why do others' evaluations affect self-esteem so strongly? If asked, most theorists would either echo Cooley's (1902) notion that people use social feedback to draw inferences about their characteristics or maintain that basing one's self-esteem on others' judgments reflects an unhealthy dependence on other people's opinions (Bednar et al., 1989; Deci and Ryan, 1995; May, 1983; Rogers, 1959).

Second, these perspectives share the view that the motive to enhance and protect one's self-esteem is generally beneficial, even if doing so requires people to hold self-serving [p. 145 ↓] illusions about themselves (Taylor and Brown, 1988). However, others have suggested that self-serving biases are maladaptive because misperceiving one's capabilities and goodness sets one up for behavioral miscalculations and maladjustment (Branden, 1969; Colvin and Block, 1994; Robins and Beer, 2001). If we assume that adjustment and success require people to estimate their attributes and abilities reasonably accurately, overestimating one's competence, making excuses, and unfairly blaming other people might be expected to undermine rather than enhance performance and wellbeing (Baumeister et al., 1993; Colvin and Block, 1994; Robins and Beer, 2001). If so, why do people have an inherent need to enhance their self-esteem?

Personal Narrative of the Development of Sociometer Theory

The idea for sociometer theory emerged amidst the questions and controversies that swirled around the construct of self-esteem in the early 1990s. Spurred partly by my reading in the emerging field of evolutionary psychology, I began to wonder about the adaptive significance of self-esteem. My starting assumption was that if self-esteem – and the self-esteem motive – were as important as psychologists seemed to think, they must serve some fundamentally important function that goes beyond merely making people feel good (or bad) about themselves. Both dominance theory (Barkow, 1980) and terror management theory (Solomon et al., 1991) had offered such arguments, but neither seemed broadly tenable. On one hand, research suggested that self-esteem was more strongly linked to being liked rather than being dominant, and on the other, I could not understand how evolutionary processes could have created a self-esteem system that lowered an organism's fear of death as terror management theory suggested. I also found it questionable that the processes identified by terror management theory, which are firmly rooted in culture, could have occurred until the appearance of culture, a mere 40–60,000 years ago (Leary, 2002, 2004).

About this time, I was also becoming interested in the emotional and behavioral effects of interpersonal acceptance and rejection. Much of my earlier research had dealt with people's concerns with how they were perceived and evaluated by others, but in the early 1990s, I began to realize that one fundamental concern underlying people's desire to be evaluated positively was their desire to be accepted and to belong to groups. About this time, I contributed a brief invited response to an article in which Baumeister and Tice (1990) proposed that a great deal of anxiety arises from people's concerns with belonging and acceptance. In my comment (Leary, 1990), I suggested that although Baumeister and Tice were probably correct, many other social emotions – including jealousy, loneliness, and depression – were also reactions to rejection.

Furthermore, in reviewing the literature on the relationship between interpersonal rejection and emotion, I found that social exclusion consistently correlated not only with negative emotions but also with lower self-esteem and that events that evoke negative

social emotions also tend to lower self-esteem. In discussing this fact, I proposed the idea that self-esteem may be a

reflection of the individual's assessment of the implications of his or her behavior for social inclusion and exclusion. "State" self-esteem is tied to one's assessment of inclusion in the immediate situation; "trait" self-esteem is a compilation of the individual's history of experienced inclusion and exclusion. (Leary, 1990: 227)

In addition, I suggested that "people behave in ways that maintain self-esteem, not because of a need to preserve self-esteem per se but because such behaviors decrease the likelihood that they will be ignored, avoided, or rejected" and that behaviors that have been conceptualized as efforts to maintain self-esteem – such as approval seeking, [p. 146 ↓] self-handicapping, and self-serving attributions – are ways of maintaining or improving social acceptance (Leary, 1990: 227).

These two speculations, which at the time had only indirect empirical support, became the cornerstones of sociometer theory. As my students and I began to work on these ideas, the late Michael Kernis invited me to contribute a chapter to an edited book on self-esteem, and I took that opportunity to lay out the basic ideas of sociometer theory in a venue in which I was less fettered by the lack of empirical evidence than I would be in a peer-reviewed article (Leary and Downs, 1995). About the same time, my students and I published a five-study article that offered the first direct tests of the theory (Leary et al., 1995).

Sociometer Theory: Details

The central premise of sociometer theory is that self-esteem is a subjective gauge of the degree to which people perceive that they are relationally valued and socially accepted by other people. State self-esteem – people's current feelings about themselves that fluctuate in response to situational events – is conceptualized as a reaction to people's perceptions of the degree to which they are, or are likely to be, valued and accepted by other people in the immediate context or near future. Trait self-esteem – people's average or typical level of self-esteem across situations and time – is conceptualized

as a reflection of their sense of the degree to which they are generally valued and accepted by others. Sociometer theory differs from most early explanations of self-esteem in suggesting that self-esteem has no value in its own right and that people neither need self-esteem nor are motivated to pursue it for its own sake. Rather, self-esteem is viewed as the output of a system that monitors and responds to events vis-à-vis interpersonal acceptance and rejection.

Although the evolutionary underpinnings were not a necessary part of the theory, we speculated regarding why human beings might have evolved a system for monitoring their relational value and acceptance. Many biologists, ethologists, and psychologists have suggested that human beings and their hominid ancestors were able to survive only because they lived in cooperative groups (Ainsworth, 1989; Kameda and Tindale, 2006). Without the natural defenses of most terrestrial animals (such as fangs, claws, strength, and the ability to run fast), solitary hominids could not have survived. By living in cooperative groups, however, they could fare reasonably well. Previous theorists had noted that people have a strong motivation toward affiliation and group living, but these sociopetal motives tell only half of the story. To survive and reproduce in nature, people not only must desire other people's company but also must assure that they are accepted rather than rejected by others (Baumeister and Leary, 1995). As a result, natural selection may have favored individuals who attended to the degree to which others accepted them and behaved in ways that led others to value, accept, and support them and refrained from doing things that might lead to rejection and expulsion. Given that social acceptance was literally vital, we hypothesized that a psychological system evolved that monitored and responded to cues indicating that the person may be devalued and rejected by other people.

In our early descriptions of the theory, we described self-esteem as a gauge of the degree to which the person is "being included versus excluded by other people" or an index of their "inclusionary status" (Leary et al., 1995: 519). However, we later began to refer to self-esteem as a marker of one's *relational value* to other people. Relational value refers to the degree to which a person regards his or her relationship with another individual as valuable or important. The higher a person's relational value to other people, the more likely they are to include and support him or her (Leary, 2001). We changed our terminology because concepts such as inclusion/exclusion and acceptance/rejection connoted dichotomous states, whereas people subjectively

[p. 147 ↓] experience degrees of acceptance and rejection. Furthermore, people may feel rejected even when they have not, strictly speaking, been excluded or rejected. For example, a woman who knows that her husband loves her may nonetheless feel rejected, experience hurt feelings, and show lowered self-esteem if he chooses to watch sports on television rather than to have dinner with her. These reactions occur not because she is excluded or rejected in an objective sense but because, at that moment, she perceives that her relational value is lower than she desires (Leary, 2006).

Thus, according to sociometer theory, the common feature of events that lower self-esteem is that they potentially lower people's relational value to others. From this perspective, failure, rejection, embarrassing situations, negative evaluations, criticism, and being outperformed by other people lower self-esteem not because they damage the person's private self-image (although that may happen as well) but rather because they lower one's relational value and the probability of acceptance. Importantly, people may infer low relational value or acceptance from cues that, in themselves, do not explicitly indicate rejection. For example, being treated unfairly can lead people to feel inadequately valued (Tyler and Lind, 1992) because others typically treat those whose relationships they value fairly. Thus, people respond to injustice as if they were rejected, showing a decrease in self-esteem (Koper et al., 1993; Shroth and Shah, 2000).

The Self-Esteem Motive

People's inclination to behave in ways that maintain, protect, and increase their self-esteem has been well documented (Blaine and Crocker, 1993; Greenwald, 1980; Taylor and Brown, 1988) and, as noted, many psychologists have assumed that such actions arise from a basic need for self-esteem. Sociometer theory disputes the existence of a free-standing need for self-esteem. Rather, according to the theory, when people do things that appear intended to maintain or raise self-esteem, their goal is usually to protect and enhance their relational value and, thus, increase the likelihood of acceptance (Leary and Baumeister, 2000). The phenomena that have been linked to self-esteem maintenance – such as self-serving attributions, self-handicapping, self-promoting prejudices, and ego-defensive reactions – may reflect efforts to promote acceptance rather than to raise self-esteem per se (Leary, 2006). Granted, such actions

may boost self-esteem, but that is because those actions promote one's relational value.

Of course, people sometimes engage in cognitive tactics that help them feel better about themselves even when those tactics do not increase their relational value. However, this observation does not contradict sociometer theory's claim that the fundamental function of the self-esteem system is to monitor and respond to threats to relational value rather than to make people feel good about themselves. Nor does it suggest that people have a 'need' for self-esteem. Rather, people can sometimes think themselves into particular emotional states, using thoughts to bypass age-old monitoring systems.

Evidence

Three types of evidence support the basic propositions of sociometer theory: (1) acceptance and rejection influence state self-esteem in both laboratory experiments and real-world settings, (2) state self-esteem relates strongly to perceived social acceptance, and (3) trait self-esteem reflects people's perceptions of their general acceptability or relational value.

Effects of Acceptance and Rejection on Self-Esteem

The central prediction of sociometer theory is that acceptance and rejection influence state [p. 148 ↓] self-esteem. Data relevant to this hypothesis has been obtained in lab experiments, studies of naturally occurring rejections, and studies of changes in self-esteem over time.

Experimental Evidence

In the first test of this hypothesis (Leary et al., 1995), we wanted to demonstrate that acceptance and rejection affect state self-esteem and also that this effect is due to being rejected rather than merely “left out.” Participants came to the laboratory in groups of five and completed self-descriptive questionnaires that were ostensibly shared with the other four participants. After seeing each other's answers, each participant listed the two participants with whom he or she would most like to work on an upcoming task. Participants then received bogus feedback indicating that they had been assigned either to work with the three-person group or to work alone, and that this assignment was based on the preferences of the other members or a random procedure. As hypothesized, the results showed that not being chosen for the group significantly lowered state self-esteem, whereas being excluded for a random reason had no effect. Study 4 of Leary et al. (1995) replicated this finding using feedback from an individual rather than from a group. Participants who were led to believe that another person liked and wanted to interact with them rated themselves more positively than those who thought another person did not like or want to interact with them.

Many other studies have shown that rejection lowers self-esteem using a variety of methods of leading participants to believe that they were rejected, ostracized, ignored, or left out (Buckley et al., 2004; Leary et al., 1998; Nezlek et al., 1997; Smith and Williams, 2004; Snapp and Leary, 2001; Wilcox and Mitchell, 1977; Zadro et al., 2005). Blackhart et al. (2009) conducted a meta-analysis of over 70 experiments that examined the effects of rejection on self-esteem. Although not every study showed the effect, overall, rejection resulted in lower self-esteem than acceptance, with an average effect size of 0.30.

Leary et al. (1998) conducted four experiments that examined how state self-esteem responds over a wider range of feedback than mere acceptance or rejection. As predicted, negative, rejecting feedback consistently lowered self-esteem relative to positive, accepting feedback, but the data also revealed a significant curvilinear pattern. State self-esteem was consistently low when feedback from others was rejecting or neutral, then rose with increasing acceptance until a moderately high level of acceptance was achieved, at which point further increases in acceptance had no further

effect. This pattern shows that state self-esteem is most responsive to small variations in relational value when feedback is in the range of neutral to moderately-accepting, possibly because slight changes in relational value have the greatest potential to affect how people are treated by others within this range. Because people are treated differently depending on whether they are slightly, moderately, or highly valued, state self-esteem shows concomitant changes with varying relational value. However, once relational value becomes slightly negative, people are simply ignored or avoided, and greater negativity typically results in no worse outcomes for the rejected person than mildly negative feedback. Likewise, once relational value is relatively high, further increases have little tangible effect, so people's self-esteem does not increase further in the upper range. The patterns suggest that the sociometer distinguishes easily between clearcut rejection and clearcut acceptance but is most sensitive to slight gradations in interpersonal appraisals within the range in which such differences in relational value are most consequential.

Many people maintain that their own self-esteem is immune from the social influences that sociometer theory identifies, asserting that they evaluate themselves solely in accord with their personal standards and are unaffected by other people's judgments. To examine the validity of these claims, Leary et al. (2003a) conducted two experiments that compared the effects of rejection on people [p. 149 ↓] who do and do not believe that their self-esteem is affected by other people's evaluations. In both studies, two extreme groups of participants were preselected – one composed of people who strongly and unequivocally maintained that their self-esteem is unaffected by disapproval and rejection, and one composed of people who strongly and unequivocally admitted that their self-esteem is affected by disapproval and rejection. Participants received accepting or rejecting feedback, and their state self-esteem was assessed. Results showed that rejection affected participants' self-esteem equally whether or not they believed that others' evaluations would influence their self-esteem. People who deny that their self-feelings are affected by acceptance and rejection are simply wrong.

The theory predicts that self-esteem is linked to traits that promote relational value and acceptance from other people. Because the characteristics that promote relational value differ across relationships, self-esteem may be attuned to different characteristics in different contexts. Across several studies, Anthony et al. (2007a) showed that self-esteem is generally attuned more strongly to social commodities such as physical

attractiveness, popularity, and social skills than to communal qualities such as warmth, kindness, and honesty, possibly because social commodities are more easily observed by other people, and acceptance in many contexts depends more strongly on social commodities than communal qualities. However, in close ongoing relationships and intimate contexts, communal characteristics become more important indicators of people's relational value. Anthony et al. showed that the sociometers of people in romantic relationships are more attuned to their communal qualities than are the sociometers of people who are not in relationships. When social roles emphasize communal qualities for social acceptance, people's self-esteem is more tightly related to their beliefs about their kindness, supportiveness, and honesty.

Correlational Evidence

Studies also show that rejection episodes lower self-esteem in people's everyday lives (Leary et al., 1998; Sommer et al., 2001; Williams et al., 1998). In their analyses of real-world cases of ostracism, Williams and his colleagues have repeatedly found that ostracism lowers self-esteem (Williams, 2001; Williams et al., 1998; Williams and Zadro, 2001). Indeed, Williams (2001) suggested that self-esteem is one of four basic needs threatened by ostracism. (In denying the existence of a need for self-esteem, sociometer theory interprets the data differently, suggesting instead that ostracism affects the sociometer and thus self-esteem, but the point remains that ostracism lowers self-esteem.) Similarly, rejecting events that hurt people's feelings also lower their self-esteem, and the magnitude of the change in self-esteem correlates with how rejected people feel (Leary et al., 1998).

In a study of married people, Shackelford (2001) found that self-esteem was inversely related to the degree to which people's spouses engaged in behaviors that inflict costs, such as infidelity, acting in a condescending manner, withholding sex, and derogating their spouse's physical attractiveness. Shackelford interpreted these findings as evidence that self-esteem tracks "spousal cost-infliction," but the results are consistent with sociometer theory. People whose spouses behave in these ways are unlikely to feel sufficiently valued and accepted as marriage partners.

Longitudinal Evidence

Four studies have examined the impact of changes in perceived acceptance on self-esteem over time. In a daily diary study of the relationship between people's self-perceived social skills and the quality of their daily interactions, Nezlek (2001) tested 83 participants during two two-week periods separated by two years. Changes in the perceived quality of one's interactions over time predicted changes in self-perceived social skill. If we assume, as Nezlek suggested, that interaction [p. 150 ↓] quality (e.g., closeness, responsiveness) is strongly linked to perceived acceptance and that self-ratings of social skill relate strongly to self-esteem, these data suggest that positive, accepting interactions increase people's self-evaluations and self-esteem. In another diary study, Denissen et al. (2008) examined the relationship between ratings of the quantity and quality of people's social interactions and their daily self-esteem. Results showed that the quality (closeness) of people's interactions, but not the quantity of interactions, predicted self-esteem both concurrently and on the subsequent day.

In a more controlled test of the effects of others' regard on self-esteem, Srivastava and Beer (2005) had participants meet four times to work together on tasks and hold group discussions. At the end of each session, measures were taken of self-esteem, perceived regard from the other members, and liking for other members, allowing the researchers to examine the effects of members' evaluations on other participants' self-esteem. Consistent with sociometer theory predictions, both actual liking and perceived regard predicted participants' self-esteem a week later. (Reciprocal effects of self-esteem on later liking were not obtained.)

Murray et al. (2003) examined daily changes in perceived regard, concerns with acceptance, and self-esteem in 154 married or cohabiting couples over a three-week period. Among other findings, they showed that anxiety about acceptance by one's partner predicted decrements in self-esteem on the subsequent day. In brief, dozens of experimental, correlational, and longitudinal studies have provided support for the idea that changes in relational value or acceptance/rejection are associated with systematic changes in state self-esteem.

Concordance between Perceived Acceptance and Self-Esteem

A second hypothesis is that people's subjective feelings of acceptance and rejection should relate to their state self-esteem. As predicted, strong relationships between perceived acceptance and state self-esteem have been obtained.

Leary et al. (1995, Study 2) asked participants to recall a social situation, indicate how included versus excluded they felt in that situation, and rate how they felt about themselves at the time on indices of state self-esteem. Correlations between perceived inclusion/exclusion and state self-esteem were quite high, ranging from -0.68 to -0.92 (depending on which of four types of situations participants were asked to recall). Based on the strength of these findings, Leary et al. concluded that "for all practical purposes, self-feelings were a proxy for perceived exclusion" (1995: 523).

If state self-esteem monitors acceptance-rejection, then the effects of particular events on people's state self-esteem ought to mirror their assumptions about the effects of those events on social acceptance and rejection. To test this idea, Leary et al. (1995; Study 1) had participants imagine performing various behaviors that varied in social desirability (such as *I gave a dollar to a beggar*, *I volunteered to donate blood*, and *I cheated on a final exam*). After imagining each situation, participants rated both how they thought others would react toward them if they performed each behavior (1 = *many other people would reject or avoid me*; 5 = *many other people would accept or include me*) and how they would personally feel about themselves. The two sets of ratings correlated highly and the rank-order of the ratings was nearly identical. As predicted by the theory, people's feelings about themselves mirrored how they thought others would respond to the events vis-à-vis social acceptance and rejection.

Additional evidence linking perceived acceptance or relational value to state self-esteem is provided by some of the experiments described earlier. In many studies that have experimentally led participants to feel accepted or rejected, perceived acceptance-rejection (typically measured as [p. 151 ↓] a manipulation check) correlated highly with participants' state self-esteem.

Trait Self-Esteem

Although sociometer theory focuses primarily on the role of state self-esteem in monitoring and managing people's reactions to events that have implications for their relational value and acceptance, it also speaks to the nature of individual differences in trait self-esteem. In my commentary on the Baumeister and Tice (1990) article discussed earlier, I had suggested that trait self-esteem should relate to the degree to which people generally feel included versus excluded by other people as well as to their personal history of experienced inclusion and exclusion (Leary, 1990). Whereas state self-esteem monitors one's relational value at the moment, trait self-esteem tracks people's general acceptability over time.

Having a global perspective on one's overall relational value is important because how people respond to a particular interpersonal event should be influenced, in part, by their views of their long-term prospects for relationships and group memberships. Elsewhere, I explained the relationship between state and trait with an analogy to the stock market: "Just as a savvy investor must monitor both the current price and long-term prospects for a stock, people must monitor both short-term fluctuations in their relational value (state self-esteem) and their relational value in the long run (trait self-esteem)," (Leary and MacDonald, 2003: 404).

Evidence strongly supports the hypothesis that trait self-esteem reflects people's perceived relational value (for a review, see Leary and MacDonald, 2003). Trait self-esteem correlates highly with people's perceptions of the degree to which they are valued, supported, and accepted by other people (Lakey et al., 1994; Leary et al., 1995, 2001) as well as with the perceived quality of their interactions with other people (Denissen et al., 2008). Along the same lines, people who have higher self-esteem have greater confidence than people with low self-esteem that their romantic partner loves them and regards them favorably (Murray and Holmes, 2000), and people with high self-esteem anticipate greater acceptance from new interaction partners than do lows (Anthony et al., 2007b). In a particularly interesting study that used country as the unit of analysis, data from 31 countries revealed that the average self-esteem of respondents from a particular nation could be predicted by the average quantity and

quality of people's interactions with their friends, even after controlling for happiness, neuroticism, individualism, and gross domestic product (Denissen et al., 2008).

In his groundbreaking work on the antecedents of self-esteem, Coopersmith (1967) identified four primary domains from which people derive their self-esteem: significance (social attention, acceptance, and affection), adherence to moral and ethical standards, the ability to influence other people, and competence. Although Coopersmith did not appear to recognize the inherently social nature of these four dimensions or the fact that they are primary determinants of acceptance and rejection by other people, he did observe that people may develop high self-esteem by being successful in just one or two domains, as long as they are approved of by the person's primary reference group, a finding consistent with sociometer theory.

In a more nuanced examination of the relationship between dimensions of self-perceptions and trait self-esteem, MacDonald et al. (2003) showed that people's self-evaluations on particular dimensions predicted trait self-esteem primarily to the degree that people thought those dimensions were important for social approval. Participants completed a measure of trait self-esteem and rated themselves on characteristics such as competence, attractiveness, and sociability, as well as the degree to which they thought people are more approving and accepting of those who are high in each domain and the degree to which people are disapproving and rejecting of those who do not possess these attributes [p. 152 ↓] (i.e., those who are incompetent, unattractive, and so on). Consistent with previous research, participants who believed that they possessed more socially desirable characteristics (i.e., characteristics that promote relational value) had higher trait self-esteem. More importantly, for four of the five domains, participants with the highest trait self-esteem were people who both evaluated themselves highly in that domain and also indicated that the domain had implications for acceptance and rejection. For example, self-evaluations of competence, attractiveness, and material wealth predicted trait self-esteem more strongly when participants believed that these domains led to approval, whereas participants who thought great material wealth leads to rejection had lower self-esteem the wealthier they were! Thus, as sociometer theory predicts, believing that one possesses certain desirable characteristics predicts trait self-esteem primarily to the degree that people think that the characteristic leads to acceptance or forestalls rejection. Similar findings

show that people whose peer group disdains success show increased self-esteem when they fail (Jones et al., 1990).

Social and Psychological Problems

Psychology's interest in self-esteem is based partly on the assumption that low self-esteem is associated with certain emotional, behavioral, and societal problems. The relationship between low self-esteem and dysfunctional behavior is not as strong as many suppose (Baumeister et al., 2003; Mecca et al., 1989), but in general, people with low self-esteem tend to show greater depression, anxiety, alcohol and drug abuse, educational and occupational difficulties, conflicted relationships, and other patterns of problem behaviors than people with higher self-esteem (see Leary and MacDonald, 2003; Leary et al., 1995; Mruk, 1995). Because they have assumed that low self-esteem causes these kinds of difficulties, many psychologists and social engineers have advocated raising self-esteem as a way to reduce them (Mecca et al., 1989; Mruk, 1995).

From the standpoint of sociometer theory, low self-esteem does not cause these problems, as is typically assumed, but rather is related to dysfunctional behaviors via three general routes. First, being inadequately accepted causes a number of aversive emotions (such as sadness, anger, and loneliness), as well as maladaptive interpersonal behaviors such as defensiveness, aggression, and negative judgments of other people. Not only are people with a history of rejection often depressed, anxious, hostile, and aggressive (Kelly, 2001; Leary et al., 2001, 2003b), but such reactions have also been demonstrated in response to short-term rejections in laboratory experiments (Buckley et al., 2003; Leary et al., 1995; Nezlek et al., 1997; Twenge et al., 2001). As a monitor of relational value, self-esteem decreases with perceived rejection and, thus, correlates with other emotional and behavioral reactions to rejection, but low self-esteem does not cause these problems.

Second, feeling inadequately valued typically increases people's desire to be accepted. (Baumeister and Leary, 1995). People usually prefer to build their relational value through socially desirable means such as achievement, being a nice person, or having socially desirable characteristics. However, when people perceive that these

routes to acceptance are not available, they may seek acceptance via deviant or antisocial behaviors such as drug use or gang membership. As a result, people who feel inadequately valued and, perhaps, inherently unacceptable (and who thus have lower self-esteem) may join deviant groups in which the standards for acceptance are lower than in socially acceptable groups (Haupt and Leary, 1997). Low self-esteem may correlate with deviant behavior, in part, because deviancy is one route to acceptance for those who feel unaccepted by mainstream groups.

Third, most psychological and interpersonal problems lead other people to devalue [p. 153 ↓] and distance themselves from the individual (Corrigan and Penn, 1999), thereby lowering his or her self-esteem. Thus, people who have psychological difficulties or behave in socially unacceptable ways tend to have lower self-esteem. For example, substance abuse, domestic violence, and chronic failure may lead to relational devaluation and, thus, lower self-esteem. In general, low trait self-esteem is more likely to be an effect than a cause of dysfunctional behavior (Baumeister et al., 2003).

In reconceptualizing the relationship between self-esteem and behavior, sociometer theory offers new insights regarding clinical and societal interventions for problems in which low self-esteem is implicated. From the sociometer perspective, low self-esteem is not necessarily a sign of psychological problems and, in fact, may indicate that one's sociometer is functioning properly. People who have low self-esteem because they have been rejected for misdeeds that hurt other people may be quite well adjusted; people shouldn't feel good about themselves when they act in socially undesirable ways. Of course, people are sometimes devalued or rejected for reasons that are not their fault, but even then, the sociometer should naturally respond with lowered self-esteem until it is consciously overridden by a rational analysis of the situation. Helping people to see that low self-esteem does not necessarily indicate that anything is wrong with them may be an important first step in these cases.

From the standpoint of sociometer theory, the therapeutic goal should not be to increase self-esteem (as clinical and social interventions sometimes do; see Bednar et al., 1989; Mecca et al., 1989; Mruk, 1995), because low self-esteem is regarded as a side effect or symptom of low perceived relational value rather as a problem in its own right. Rather, a sociometer-based intervention recognizes that the crux of problems associated with low self-esteem involves concerns with acceptance and rejection.

Thus, the first step in any such intervention is to determine whether the problem under consideration arises from feeling inadequately valued, reflects dysfunctional ways of seeking social acceptance, and/or leads to devaluation and rejection. Then, interventions should focus on addressing the actual source of the problem, which involves low perceived relational value. As it turns out, many clinical psychologists and social engineers inadvertently adopt this approach in that many interventions designed to raise self-esteem rely on programs that promote people's perceptions of their social desirability and relational value (Leary, 1999). But, sociometer theory focuses attention on the real genesis of most problems related to low self-esteem – inadequate relational value.

Questions and Criticisms

In this concluding section, I examine some recurring questions about sociometer theory.

The Inherently Interpersonal Nature of Self-Esteem

Perhaps the most common resistance to sociometer theory comes from those who insist that true or healthy self-esteem is, by definition, based solely on adherence to one's own personal standards and should not be affected by other people's evaluations (Deci and Ryan, 1995; May, 1983). This objection is intriguing because it implies that being concerned with other people's impressions and judgments is inherently undesirable, if not maladaptive. Yet, congenial, cooperative, and supportive interpersonal relationships, as well as people's social wellbeing, require that they pay attention to how they are being perceived, evaluated, and accepted by other people. In fact, if people have a need to belong that evolved because it conferred adaptive advantages (Baumeister and Leary, 1995), then we should expect well-functioning people to keep one eye on how other people react to them. Sociometer theory suggests that healthy [p. 154 ↓] self-esteem is responsive to the evaluations of other people, at least within bounds.

Far from being a sign of insecurity, manipulation, or vanity, being attuned to other people's reactions – and particularly to the degree to which they are accepting versus rejecting – is essential for personal and social wellbeing. Even though people can be overly concerned with others' reactions in ways that create unnecessary distress or lead them to behave in ways that hurt themselves or others, possessing a mechanism that monitors other people's reactions and affects people's self-relevant feelings when unfavorable interpersonal circumstances are detected is critically beneficial both for the person and for those individuals and groups with whom he or she interacts. In fact, maintaining close relationships and cooperative groups requires that people attend to one another's evaluations, immediately recognize instances in which their own behavior may lead others to devalue their relationship, feel badly when their actions increase social distance, and promote social connections within the relationship or group. This is not to say that people should not behave in ways that might jeopardize their relationship or group when the need arises, but as a rule, the sociometer helps people to maintain the kinds of relationships upon which cooperation and social support is based.

Other Sources of Self-Esteem

A second question is whether factors other than relational value also influence self-esteem. Even if one acknowledges that people feel better about themselves when they are accepted than rejected, one could still argue that other outcomes – such as mastering tasks, professional success, accomplishing difficult goals, or behaving in a morally exemplary manner – also increase self-esteem.

Proponents of sociometer theory concur that such actions influence self-esteem but interpret such patterns as reflecting the effects of excellence, achievement, and moral behavior on relational value and acceptance. People are more likely to be valued and accepted when they are competent, successful, and ethical than when they are not, so any action or outcome that potentially raises or lowers relational value should influence self-esteem. Even if a person's actions or outcomes are not currently known by other people, the sociometer should alert the individual via changes in state self-esteem when private actions or outcomes have potential implications for relational value (Leary and Baumeister, 2000).

Seeking Self-Esteem

Several writers have misinterpreted sociometer theory to suggest that people should seek self-esteem through social acceptance. Not only have many theorists suggested that healthy self-esteem is decidedly not tied to social approval (Deci and Ryan, 1995; May, 1983), but evidence suggests that trying to build self-esteem by seeking approval and acceptance is associated with negative affect, self-regulatory problems, defensiveness, poor interpersonal relationships, and lower mental and physical health (Crocker and Knight, 2005; Deci and Ryan, 1995; Schimel et al., 2001). Proponents of sociometer theory agree. If self-esteem is merely a gauge, people should not try to influence their self-esteem directly by any means.

Cross-Cultural Differences

Research showing that Japanese respondents do not engage in the self-enhancing biases that are seen in American and European samples has led some researchers to conclude that people in certain cultures are either indifferent to self-esteem or do not experience self-esteem at all (Heine et al., 1999, 2000, 2001). Sociometer theory offers a different interpretation of such findings. Because the criteria for being relationally valued differ across cultures, people tend to behave [p. 155 ↓] in ways that are associated with social acceptance and refrain from behaving in ways that lead to rejection in their culture. Unlike in the US, where maintaining a confident, self-enhancing persona is acceptable, in Japan, self-enhancement does not promote acceptance and, in fact, may result in disapproval and devaluation. As a result, the Japanese are less likely to self-enhance and feel more badly about themselves when they do. Research suggests that behaviors that promote relational value in Japan – self-effacement and deference, for example – result in high self-esteem and that the Japanese “enhance” on communal, self-effacing attributes (Sedikides et al., 2003).

How Many Sociometers?

We initially assumed that people possess a single sociometer that monitors their relational value across a wide array of relationships and groups. However, based on evolutionary logic, Kirkpatrick and Ellis (2001) suggested that people may possess a number of domain-specific sociometers that monitor relational value in qualitatively different kinds of relationships. They noted that the criteria for acceptance into and rejection from different kinds of relationships – such as family relationships, instrumental coalitions, and mating relationships – differ greatly, as do the means of increasing one's acceptance in each of these kinds of relationships. Given that evolutionary processes would be unlikely to create a single, multipurpose sociometer to manage relational value across all kinds of relationships, human beings may have evolved specific sociometers for dealing with interpersonal problems in different social domains. Kirkpatrick and Ellis also suggested that these sociometers may serve functions in addition to promoting social acceptance, such as facilitating personality development, addressing personal deficiencies, and guiding adaptive relationship choices (Kirkpatrick and Ellis, 2001, 2006).

Conclusion

Whatever verdict the field reaches regarding the merits of sociometer theory as an explanation of self-esteem, the ultimate contribution of the sociometer approach may lie as much in the way that it has recast longstanding questions about self-esteem as in the particulars of the theory itself. Sociometer theory has forced researchers to consider the functions of self-esteem and particularly to consider functions of self-esteem that go beyond merely defending the ego or reducing anxiety. It has also helped to focus attention on fundamental social psychological processes involved in how people respond to interpersonal rejection.

In addition, sociometer theory has led many psychologists to question the assumption that people have a motive or need for self-esteem. When examined critically, one finds it difficult to identify precisely what purpose a need to feel good about oneself would

serve unless it were associated with important tangible outcomes. Even so, many writers continue to explain various behaviors by invoking the idea that people are trying to protect or enhance their self-esteem. In almost all instances, sociometer theory offers a more parsimonious explanation of the effects of both self-esteem and self-esteem threats in terms of people's efforts to maintain a minimum degree of social acceptance. The idea that people have a need for self-esteem should be retired, and this should be done even if sociometer theory should eventually be found to be lacking as an explanation of self-esteem. Surely, a psychological process that is as pervasive and as related to important outcomes as self-esteem does more than simply make people feel good about themselves.

References

Ainsworth, M.S. Attachments beyond infancy . American Psychologist , 44, 709–716. (1989) <http://dx.doi.org/10.1037/0003-066X.44.4.709>

Anthony, D.B., Holmes, J.G. and Wood, J.V. Social acceptance and self-esteem: Tuning the sociometer to interpersonal value . Journal of Personality and Social Psychology , 92, 1024–1039. (2007a) <http://dx.doi.org/10.1037/0022-3514.92.6.1024>

Anthony, D.B., Wood, J.V. and Holmes, J.G. Testing sociometer theory: Self-esteem and the importance of acceptance for social decision-making . Journal of Experimental Social Psychology , 43, 425–432. (2007b) <http://dx.doi.org/10.1016/j.jesp.2006.03.002>

Bandura, A. Self-efficacy: Toward a unifying theory of behavioral change . Psychological Review , 84, 191–215. (1977) <http://dx.doi.org/10.1037/0033-295X.84.2.191>

Barkow, J.H. (1980) Prestige and self-esteem: A biosocial interpretation . In D.R. Omars, ed. , F. Strayer, ed. and D.G. Freedman (eds), Dominance Relations: An Ethological View of Human Conflict and Social Interaction , pp. 319–332. New York: Garland STPM.

Baumeister, R.F., Campbell, J.D., Krueger, J.I. and Vohs, K.D. Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest*, 4, 1–44. (2003) <http://dx.doi.org/10.1111/1529-1006.01431>

Baumeister, R.F., Heatherton, T.F. and Tice, D.M. When ego threats lead to self-regulation failure: Negative consequences of high self-esteem . *Journal of Personality and Social Psychology* , 64, 141–156. (1993) <http://dx.doi.org/10.1037/0022-3514.64.1.141>

Baumeister, R.F. and Leary, M.R. The need to belong: Desire for interpersonal attachments as a fundamental human motivation . *Psychological Bulletin* , 117, 497–529. (1995) <http://dx.doi.org/10.1037/0033-2909.117.3.497>

Baumeister, R.F. and Tice, D.M. Anxiety and social exclusion . *Journal of Social and Clinical Psychology* , 9, 165–195. (1990) <http://dx.doi.org/10.1521/jscp.1990.9.2.165>

Bednar, R.L., Wells, M.G. and Peterson, S.R. (1989) Self-esteem: Paradoxes and Innovations in Clinical Theory and Practice . Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10068-000>

Blackhart, G.C., Nelson, B.C., Knowles, M.L. and Baumeister, R.F. Rejection elicits emotional reactions but neither causes immediate distress nor lowers self-esteem: A meta-analytic review of 192 studies on social exclusion . *Personality and Social Psychology Review* , 13, 269–309. (2009) <http://dx.doi.org/10.1177/1088868309346065>

Blaine, B. and Crocker, J. (1993) Self-esteem and self-serving biases in reactions to positive and negative events: An integrative review . In R.F. Baumeister (ed.), *Self-esteem: The Puzzle of Low Self-Regard* , pp. 55–85. New York: Plenum.

Bowles, T. Focusing on time orientation to explain adolescent self concept and academic achievement: Part II. Testing a model . *Journal of Applied Health Behaviour* , 1, 1–8. (1999)

Branden, N. (1969) *The Psychology of Self-esteem* . New York: Bantam.

- Branden, N. (1983) Honoring the Self . Los Angeles. Bantam Books.
- Buckley, K.E., Winkel, R.E. and Leary, M.R. Reactions to acceptance and rejection: Effects of level and sequence of relational evaluation . Journal of Experimental Social Psychology , 40, 14–28. (2004) <http://dx.doi.org/10.1016/S0022-1031%2803%2900064-7>
- Cervone, D., Artisico, D. and Berry, J.M. (2006) Self-efficacy and adult development . In C. Hoare (ed.), Handbook of Adult Development and Learning , pp. 169–195. New York: Oxford University Press.
- Colvin, C.R. and Block, J. Do positive illusions foster mental health? An examination of the Taylor and Brown formulation . Psychological Bulletin , 116, 3–20. (1994) <http://dx.doi.org/10.1037/0033-2909.116.1.3>
- Cooley, C.H. (1902). Human Nature and the Social Order . New York: Scribner.
- Coopersmith, S. (1967) Antecedents of Self-esteem . San Francisco: Freeman.
- Corrigan, P.W. and Penn, D.L. Lessons from social psychology on discrediting psychiatric stigma . American Psychologist , 54, 765–776. (1999) <http://dx.doi.org/10.1037/0003-066X.54.9.765>
- Crocker, J. and Knight, K.M. Contingencies of self-worth . Current Directions in Psychological Science , 14, 200–203. (2005) <http://dx.doi.org/10.1111/j.0963-7214.2005.00364.x>
- Deci, E.L. and Ryan, R.M. (1995) Human agency: The basis for true self-esteem . In M.H. Kernis (ed.), Efficacy, Agency, and Self-esteem , pp. 31–50. New York: Plenum.
- Denissen, J.J.A., Penke, L., Schmitt, D.P. and van Aken, M.A.G. Self-esteem reactions to social interactions: Evidence for sociometer mechanisms across days, people, and nations . Journal of Personality and Social Psychology , 95, 181–196. (2008) <http://dx.doi.org/10.1037/0022-3514.95.1.181>

Glendinning, A. and Inglis, D. Smoking behaviour in youth: The problem of low self-esteem? *Journal of Adolescence*, 22, 673–682. (1999) <http://dx.doi.org/10.1006/jado.1999.0262>

Greenwald, A.G. The totalitarian ego: Fabrication and revision of personal history. *American Psychologist*, 35, 603–613. (1980) <http://dx.doi.org/10.1037/0003-066X.35.7.603>

Guay, F., Delisle, M., Fernet, C., Julien, É. and Senécal, C. Does task-related identified regulation moderate the sociometer effect? A study of performance feedback, perceived inclusion, and state self-esteem. *Social Behavior and Personality*, 36, 239–254. (2008) <http://dx.doi.org/10.2224/sbp.2008.36.2.239>

Hansford, B.C. and Hattie, J.A. The relationship between self and achievement/performance measures. *Review of Educational Research*, 52, 123–142. (1982)

Haupt, A.H. and Leary, M.R. The appeal of worthless groups: Moderating effects of trait self-esteem. *Group Dynamics: Theory, Research, and Practice*, 1, 124–132. (1997) <http://dx.doi.org/10.1037/1089-2699.1.2.124>

Heaven, P.C. Authoritarianism, directiveness, and self-esteem revisited: A cross-cultural analysis. *Personality and Individual Differences*, 7, 225–228. (1986) <http://dx.doi.org/10.1016/0191-8869%2886%2990059-0>

Heine, S.H., Lehman, D.R., Markus, H.R. and Kitayama, S. Is there a universal need for positive self-regard? *Psychological Review*, 106, 766–794. (1999) <http://dx.doi.org/10.1037/0033-295X.106.4.766>

Heine, S.J., Kitayama, S. and Lehman, D.R. Cultural differences in self-evaluation: Japanese readily accept negative self-relevant information. *Journal of Cross Cultural Psychology*, 32, 434–443. (2001) <http://dx.doi.org/10.1177/0022022101032004004>

Heine, S.J., Takata, T. and Lehman, D.R. Beyond self-presentation: Evidence for self-criticism among Japanese. *Personality and Social Psychology Bulletin*, 26, 71–78 (2000) <http://dx.doi.org/10.1177/0146167200261007>

James, W. (1890) *The Principles of Psychology* . New York: Henry Holt. <http://dx.doi.org/10.1037/11059-000>

Jones, E.E., Brenner, K.J. and Knight, J.G. When failure elevates self-esteem . *Personality and Social Psychology Bulletin* , 16, 200–209. (1990) <http://dx.doi.org/10.1177/0146167290162002>

Kameda, T. and Tindale, R.S. (2006) Groups as adaptive devices: Human docility and group aggregation mechanisms in evolutionary context . In M. Schaller, ed. , J.A. Simpson, ed. , and D.T. Kenrick (eds), *Evolution and Social Psychology* , pp. 317–341. New York: Psychology Press.

Kelly, K.M. (2001) Individual differences in reactions to rejection . In M.R. Leary (ed.), *Interpersonal Rejection* , pp. 291–315. New York: Oxford University Press.

Kirkpatrick, L.A. and Ellis, B.J. (2001) An evolutionary-psychological approach to self-esteem: multiple domains and multiple functions . In G.J.O. Fletcher, ed. and M.S. Clark (eds), *Blackwell Handbook of Social Psychology: Interpersonal Processes* , pp. 411–436. Oxford: Blackwell.

Kirkpatrick, L.A. and Ellis, B.J. (2006) The adaptive functions of self-evaluative psychological mechanisms . In M.H. Kernis (ed.), *Self-esteem: Issues and Answers* , pp. 334–339. New York: Psychology Press.

Koper, G., van Knippenberg, D., Bouhuijs, F., Vermunt, R. and Wilke, H. Procedural fairness and self-esteem . *European Journal of Social Psychology* , 23, 313–325. (1993) <http://dx.doi.org/10.1002/ejsp.2420230307>

Lakey, B., Tardiff, T.A. and Drew, J.B. Negative social interactions: Assessment and relations to social support, cognition, and psychological distress . *Journal of Social and Clinical Psychology* , 13, 42–62. (1994) <http://dx.doi.org/10.1521/jscp.1994.13.1.42>

Leary, M.R. Responses to social exclusion: Social anxiety, jealousy, loneliness, depression, and low self-esteem . *Journal of Social and Clinical Psychology* , 9, 221–229. (1990) <http://dx.doi.org/10.1521/jscp.1990.9.2.221>

- Leary, M.R. (1999) The social and psychological importance of self-esteem . In R.M. Kowalski, ed. and M.R. Leary (eds), *The Social Psychology of Emotional and Behavioral Problems: Interfaces of Social and Clinical Psychology* , pp. 197–221. Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10320-007>
- Leary, M.R. (2001) Toward a conceptualization of interpersonal rejection . In M.R. Leary (ed.), *Interpersonal Rejection* , pp. 3–20. New York: Oxford University Press.
- Leary, M.R. (2002). The interpersonal basis of self-esteem: Death, devaluation, or deference? In J. Forgas, ed. and K.D. Williams (eds), *The Social Self: Cognitive, Interpersonal, and Intergroup Perspectives* . New York: Psychology Press.
- Leary, M.R. The function of self-esteem in terror management theory and sociometer theory: A comment on Pyszczynski, Greenberg, Solomon, Arndt, and Schimel (2004) . *Psychological Bulletin* , 130, 478–482. (2004) <http://dx.doi.org/10.1037/0033-2909.130.3.478>
- Leary, M.R. (2005) Varieties of interpersonal rejection . In K.D. Williams, ed. , J.P. Forgas, ed. , and W. von Hippel (eds), *The Social Outcast* , pp. 35–51. New York: Psychology Press.
- Leary, M.R. Sociometer theory and the pursuit of relational value: Getting to the root of self-esteem . *European Review of Social Psychology* , 16, 75–111. (2006) <http://dx.doi.org/10.1080/10463280540000007>
- Leary, M.R. and Baumeister, R.F. (2000) The nature and function of self-esteem: Sociometer theory . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 32, 1–62. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2800%2980003-9>
- Leary, M.R., Cottrell, C.A. and Phillips, M. Deconfounding the effects of dominance and social acceptance on self-esteem . *Journal of Personality and Social Psychology* , 81, 898–909. (2001) <http://dx.doi.org/10.1037/0022-3514.81.5.898>

Leary, M.R. and Downs, D.L. (1995) Interpersonal functions of the self-esteem motive: The self-esteem system as a sociometer . In M. Kernis (ed.), *Efficacy, Agency, and Self-esteem* , pp. 123–144. New York: Plenum.

Leary, M.R., Gallagher, B., Fors, E.H., Buttermore, N., Baldwin, E., Lane, K.K. and Mills, A., The invalidity of disclaimers about the effects of social feedback on self-esteem . *Personality and Social Psychology Bulletin* , 29, 623–636. (2003a) <http://dx.doi.org/10.1177/0146167203029005007>

Leary, M.R., Haupt, A., Strausser, K. and Chokel, J. Calibrating the sociometer: The relationship between interpersonal appraisals and state self-esteem . *Journal of Personality and Social Psychology* , 74, 1290–1299. (1998) <http://dx.doi.org/10.1037/0022-3514.74.5.1290>

Leary, M.R., Kowalski, R.M., Smith, L. and Phillips, S. Teasing, rejection, and violence: Case studies of the school shootings . *Aggressive Behavior* , 29, 202–214. (2003b) <http://dx.doi.org/10.1002/ab.10061>

Leary, M.R. and MacDonald, G. (2003) Individual differences in self-esteem: A review and theoretical integration . In M.R. Leary, ed. and J.P. Tangney (eds), *Handbook of Self and Identity* , pp. 401–418. New York: Guilford Press.

Leary, M.R., Schreindorfer, L.S. and Haupt, A.L. The role of low self-esteem in emotional and behavioral problems: Why is low self-esteem dysfunctional? *Journal of Social and Clinical Psychology* , 14, 297–314. (1995) <http://dx.doi.org/10.1521/jscp.1995.14.3.297>

Leary, M.R., Springer, C., Negel, L., Ansell, E. and Evans, K. The causes, phenomenology, and consequences of hurt feelings . *Journal of Personality and Social Psychology* , 74, 1225–1237. (1998) <http://dx.doi.org/10.1037/0022-3514.74.5.1225>

Leary, M.R., Tambor, E.S., Terdal, S.K. and Downs, D.L. Self-esteem as an interpersonal monitor: The sociometer hypothesis . *Journal of Personality and Social Psychology* , 68, 518–530. (1995) <http://dx.doi.org/10.1037/0022-3514.68.3.518>

LePine, J.A. and Van Dyne, L. Predicting voice behavior in work groups . Journal of Applied Psychology , 83, 853–868. (1998) <http://dx.doi.org/10.1037/0021-9010.83.6.853>

MacDonald, G., Saltzman, J.L. and Leary, M.R. Social approval and trait self-esteem . Journal of Research in Personality , 37, 23–40. (2003) <http://dx.doi.org/10.1016/S0092-6566%2802%2900531-7>

May, R. (1983) The Discovery of Being . New York: Norton.

Mecca, A.M., Smelser, N.J. and Vasconcellos, J. (1989) The Social Importance of Self-esteem . Berkeley, CA: University of California Press.

MrukC. (1995) Self-esteem: Research, Theory, and Practice . New York: Springer.

Murray, S.L., Griffin, D.W., Rose, P. and Bellavia, G.M. Calibrating the sociometer: The relational contingencies of self-esteem . Journal of Personality and Social Psychology , 85, 63–84. (2003) <http://dx.doi.org/10.1037/0022-3514.85.1.63>

Murray, S.L. and Holmes, J.G. (2000) Seeing the self through a partner's eyes: Why self-doubts turn into relationship insecurities . In A. Tesser, ed. , R.B. Felson, ed. and J.M. Suls (eds), Psychological Perspectives on Self and Identity , pp. 173–198. Washington, DC: APA Publications. <http://dx.doi.org/10.1037/10357-007>

Nezlek, J.B. Causal relationships between perceived social skills and day-to-day social interaction: Extending the sociometer hypothesis . Journal of Social and Personal Relationships , 18, 386–403. (2001) <http://dx.doi.org/10.1177/0265407501183005>

Nezlek, J.B., Kowalski, R.M., Leary, M.R., Blevins, T. and Holgate, S. Personality moderators of reactions to interpersonal rejection: Depression and trait self-esteem . Personality and Social Psychology Bulletin , 23, 1235–1244. (1997) <http://dx.doi.org/10.1177/01461672972312001>

Perez, R.C. The effect of experimentally-induced failure, self-esteem, and sex on cognitive differentiation . Journal of Abnormal Psychology , 81, 74–79. (1973) <http://dx.doi.org/10.1037/h0033974>

Pyszczynski, T., Greenberg, J., Solomon, S., Arndt, J. and Schimel, J. Why do people need self-esteem? A theoretical and empirical review . *Psychology Bulletin* , 3, 435–468. (2004) <http://dx.doi.org/10.1037/0033-2909.130.3.435>

Raskin, R., Novacek, J. and Hogan, R. Narcissistic self-esteem management . *Journal of Personality and Social Psychology* , 60, 911–918. (1991) <http://dx.doi.org/10.1037/0022-3514.60.6.911>

Robins, R.W. and Beer, J.S. Positive illusions about the self: Short-term benefits and long-term costs . *Journal of Personality and Social Psychology* , 80, 340–352. (2001). <http://dx.doi.org/10.1037/0022-3514.80.2.340>

Rogers, C. (1959) A theory of therapy, personality, and interpersonal relationships, as developed in the client-centered framework . In S. Koch (ed.), *Psychology: A Study of a Science* , 3, 184–256. New York: McGraw-Hill.

Schimel, J., Arndt, J., Pyszczynski, T. and Greenberg, J. Being accepted for who we are: Evidence that social validation of the intrinsic self reduces general defensiveness . *Journal of Personality and Social Psychology* , 80, 35–52. (2001) <http://dx.doi.org/10.1037/0022-3514.80.1.35>

Sedikides, C., Gaertner, L. and Toguchi, Y. Pancultural self-enhancement . *Journal of Personality and Social Psychology* , 84, 60–70. (2003) <http://dx.doi.org/10.1037/0022-3514.84.1.60>

Shackelford, T.K. Self-esteem in marriage . *Personality and Individual Differences* , 30, 371–390. (2001) <http://dx.doi.org/10.1016/S0191-8869%2800%2900023-4>

Shraugher, J.S. and Sorman, P.B. Self-evaluations, initial success and failure, and improvement as determinants of persistence . *Journal of Consulting and Clinical Psychology* , 45, 784–795. (1977) <http://dx.doi.org/10.1037/0022-006X.45.5.784>

Shroth, H.A. and Shah, P.P. Procedures: Do we really want to know them? An examination of the effects of procedural justice on self-esteem . *Journal of Applied Psychology* , 85, 462–471. (2000) <http://dx.doi.org/10.1037/0021-9010.85.3.462>

Smith, A. and Williams, K.D. R U there? Ostracism by cell phone text messages . Group Dynamics: Theory, Research, and Practice , 8, 291–301. (2004) <http://dx.doi.org/10.1037/1089-2699.8.4.291>

Snapp, C.M. and Leary, M.R. Hurt feelings among new acquaintances: Moderating effects of interpersonal familiarity . Journal of Social and Personal Relationships , 18, 315–326. (2001) <http://dx.doi.org/10.1177/0265407501183001>

Solomon, S., Greenberg, J. and Pyszczynski, T. (1991) A terror management theory of social behavior: The psychological functions of self-esteem and cultural worldviews . In M. Zanna (ed.), Advances in Experimental Social Psychology , 24, 91–159. Orlando, FL: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960328-7>

Sommer, K.L., Williams, K.D., Ciarocco, N.J. and Baumeister, R.F. Explorations into the intrapsychic and interpersonal consequences of social ostracism . Basic and Applied Social Psychology , 23, 227–245. (2001)

Srivastava, S. and Beer, J.S. How self-evaluations relate to being liked by others: Integrating sociometer and attachment perspectives . Journal of Personality and Social Psychology , 89, 966–977. (2005) <http://dx.doi.org/10.1037/0022-3514.89.6.966>

Taylor, S.E. and Brown, J.D. Illusion and wellbeing: A social psychological perspective on mental health . Psychological Bulletin , 103, 193–210. (1988) <http://dx.doi.org/10.1037/0033-2909.103.2.193>

Tedeschi, J.T. and Norman, N. (1985) Social power, self-presentation, and the self . In B.R. Schlenker (ed.), The Self and Social Life , pp. 293–322. New York: McGraw-Hill.

Twenge, J.M., Baumeister, R.F., Tice, D.M. and Stucke, T.S. If you can't join them, beat them: Effects of social exclusion on aggressive behavior . Journal of Personality and Social Psychology , 81, 1058–1069. (2001) <http://dx.doi.org/10.1037/0022-3514.81.6.1058>

Tyler, T.R. and Lind, E.A. (1992) A relational model of authority in groups . In M. Zanna (ed.), Advances in Experimental Social Psychology , 25, 115–191. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960283-X>

Wilcox, J. and Mitchell, J. Effects of group acceptance/rejection of self-esteem levels of individual group members in a task-oriented problem-solving group interaction . *Small Group Behavior* , 8, 169–178. (1977) <http://dx.doi.org/10.1177/104649647700800204>

Williams, K.D. (2001) Ostracism: The Power of Silence . New York: Guilford Press.

Williams, K.D., Shore, W.J. and Grahe, J.E. The silent treatment: Perceptions of its behaviors and associated feelings . *Group Processes and Intergroup Relations* , 1, 117–141. (1998) <http://dx.doi.org/10.1177/1368430298012002>

Williams, K.D. and Zadro, L. (2001) Ostracism: On being ignored, excluded and rejected . In M.R. Leary (ed.), *Interpersonal Rejection* , pp. 21–53. New York: Oxford University Press.

Zadro, L., Williams, K.D. and Richardson, R. How low can you go? Ostracism by a computer lowers belonging, control, self-esteem and meaningful existence . *Journal of Experimental Social Psychology* , 40, 560–567. (2004) <http://dx.doi.org/10.1016/j.jesp.2003.11.006>

Zadro, L., Williams, K.D. and Richardson, R. Riding the ‘O’ train: Comparing the effects of ostracism and verbal dispute on targets and sources . *Group Processes and Intergroup Relations* , 8, 125–143. (2005) <http://dx.doi.org/10.1177/1368430205051062>

<http://dx.doi.org/10.4135/9781446249222.n33>

Handbook of Theories of Social Psychology

Attachment Theory

Contributors: Mario Mikulincer & Phillip R. Shaver

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Attachment Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n34>

Print pages: 160-180

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n34>

[p. 160 ↓]

Chapter 34: Attachment Theory

Mario Mikulincer Phillip R. Shaver

Abstract

In this chapter we recount the origins and development of attachment theory, explain how it has been extended into the domain of adult personality and social psychology, and describe its six major components: (1) the innate attachment behavioral system, (2) attachment-related affect-regulation strategies, (3) internal working models of self and others, (4) attachment patterns or “styles,” (5) attachment security viewed as a resilience resource, and (6) dysfunctional aspects of attachment insecurity. We review some of the empirical literature based on these constructs and explain how attachment-related strategies affect such social psychological constructs as self-esteem, affect and affect regulation, mental health, person perception, relationship functioning and satisfaction, prejudice and intergroup hostility, prosocial behavior, leadership, and group functioning. We briefly describe how attachment theory relates to some of the other theories discussed in this handbook and explain how the theory and the research it has generated are currently being, and in the future can be, applied.

Introduction

In this chapter we explain attachment theory (Bowlby, 1982), its origins, development, and operationalization in studies of human infants and their parents; and our extension of the theory to make it suitable for personality and social psychological studies of adolescents, adults, and their close interpersonal relationships (Mikulincer and Shaver, 2007a; Shaver et al., 1988). The theory deals with close relationships and their psychological foundations and consequences. It integrates insights from psychoanalytic theory; primate ethology; cognitive developmental and social cognitive psychology;

theories of stress and coping; and contemporary research on personality development, affect regulation, and relational interdependence.

The chapter begins with an account of the origins of attachment theory during the 1960s and 1970s in the work of John Bowlby and Mary Ainsworth. We then explain how the two of us became involved with extensions of the theory in the 1980s. Next, we describe the theory itself in some detail, placing special emphasis on our version of it, which began in the work of Hazan and Shaver (1987). This version of the theory is designed for and influenced by contemporary social psychologists. We show how attachment security, grounded in responsive, supportive relationships, plays a role in personal and social issues of interest to personality-social [p. 161 ↓] psychologists, issues such as self-esteem, person perception, interpersonal behavior, exploration and achievement, and prosocial behavior and intergroup relations. We then consider the darker side of attachment relationships, which includes the defenses and personal and relational difficulties that stem from attachment insecurities. We conclude with a brief discussion of connections between attachment theory and other social psychological theories and a brief consideration of applications of attachment theory and research.

The Origins of Attachment Theory

Attachment theory was originally proposed by John Bowlby, a British psychoanalyst, and was then strengthened by the theoretical, psychometric, and empirical contributions of Mary Ainsworth, an American developmental psychologist. As explained later in this chapter, Bowlby's theory and Ainsworth's findings were extended into social psychology by Hazan and Shaver (1987), culminating 20 years later in a book-length overview of the theory and the large research literature it has generated (Mikulincer and Shaver, 2007a).

John Bowlby

Bowlby was born in England, in 1907, to economically comfortable and well-educated parents. His father was a physician, and John eventually became one as well – a psychiatrist. While studying to become a child psychiatrist, Bowlby undertook

psychoanalytic training with a famous mentor, Melanie Klein, and was psychoanalyzed for several years by Joan Riviere, a close associate of Klein's. From these mentors Bowlby learned about early relationships with caregivers; the tendency of troubled children to deal with painful experiences, especially separations and losses, by defensively excluding them from conscious memory; and the emotions of anxiety, anger, and sadness. Despite absorbing many of Klein and Riviere's ideas, however, Bowlby rejected their extreme emphasis on fantasies at the expense of reality, and on sexual drives rather than other kinds of relational motives.

Attachment theory also grew out of Bowlby's experiences as a family therapist at the Tavistock Clinic in London, where social and family relationships were considered alongside individual psychodynamics as causes of psychological and social disorders. Bowlby was also influenced by preparing a report for the World Health Organization on children who were homeless following World War II.

As Bowlby's clinical observations and insights accumulated, he became increasingly interested in explaining what, in his first major statement of attachment theory, he called "the child's tie to his mother" (Bowlby, 1958). In formulating the theory, he was especially influenced by Konrad Lorenz's (1952) ideas about "imprinting" in precocial birds and the writings of other ethologists and primatologists, including the primatologist Robert Hinde (1966). These authors, along with Harry Harlow (1959), had begun to show that immature animals' ties to their mothers were not simply due to classical conditioning based on feeding, as learning theorists (and using different language, psychoanalysts) had thought. Instead, Bowlby viewed the human infant's reliance on, and emotional bond with, its mother to be the result of a fundamental instinctual behavioral system that, unlike Freud's sexual libido concept, was viewed as social-relational rather than sexual. Because Bowlby relied so heavily on animal research and on the behavioral-systems construct, he was strongly criticized by other psychoanalysts for being a "behaviorist." He nevertheless continued to view himself as a psychoanalyst and a legitimate heir to Freud, which is how he is largely viewed today.

Bowlby eventually expanded his preliminary articles about core aspects of attachment into three major books, *Attachment and Loss*, Volumes 1, 2, and 3, which are now recognized as landmarks of modern psychology, [p. 162 ↓] psychiatry, and social science. The first volume, *Attachment*, was published in 1969 and revised in 1982; the

second, *Separation: Anxiety and Anger*, was published in 1973; and the third, *Loss: Sadness and Depression*, was published in 1980. These comprehensive volumes were accompanied in 1979 by a collection of Bowlby's lectures, *The Making and Breaking of Affectional Bonds*, and were supplemented in 1988 by a book, *A Secure Base*, about applying attachment theory and research to psychotherapy. Bowlby died in 1990, having won many professional awards.

Mary Ainsworth

Bowlby's major collaborator, Mary Salter Ainsworth, was born in Ohio in 1913 and received her PhD in developmental psychology from the University of Toronto in 1939, after completing a dissertation on security and dependency that was inspired by her advisor William Blatz's security theory. In her dissertation, "An Evaluation of Adjustment Based on the Concept of Security" (1940), Ainsworth mentioned for the first time what eventually became a central part of attachment theory, the secure-base construct, which emphasized the importance in child–parent attachment relationships of parents' provision of what Ainsworth called a secure base from which to explore the world.

When she moved to London with her anthropologist husband, Ainsworth answered a newspaper ad for a research position with Bowlby, having not known about him or his work beforehand. Part of her job was to analyze films of children's separations from mother. These films convinced her of the value of behavioral observations, which were the centerpiece of her contributions to attachment research. When her husband decided in 1953 to advance his career by undertaking cultural research in Uganda, Ainsworth moved there as well and began an observational study of mothers and infants, whom she visited every two weeks for two hours of observation over a period of several months. Eventually, after returning to North America and becoming a faculty member at the Johns Hopkins University, in 1967 Ainsworth published a book entitled *Infancy in Uganda: Infant Care and the Growth of Love*.

One of the intellectually and historically significant features of Ainsworth's 1967 book was an appendix that sketched different patterns of infant attachment, which Ainsworth linked empirically with observable aspects of maternal behavior. Although these patterns were not precisely the same as the three attachment types for which

Ainsworth later became famous (called secure, anxious, and avoidant in our work; see Ainsworth et al., 1978, for the original details), some definite similarities are evident. The three main patterns of attachment delineated in the 1978 book were based on its 1967 predecessor, but they were greatly refined by intensive studies of middle-class American infants in Baltimore. In these American studies, Ainsworth and her students recorded detailed home observations during infants' first year of life and supplemented them with a new laboratory assessment procedure: the Strange Situation. Ainsworth et al.'s 1978 book explains how to code an infant's behavior with mother in the Strange Situation, and also shows how the three major forms of infant attachment are associated with particular patterns of maternal behavior at home. The measures and ideas advanced in the 1978 book, taken in conjunction with Bowlby's theoretical books on attachment, separation, and loss form the backbone of all subsequent discussions of normative attachment processes and individual differences in attachment behavior. This work provided a foundation for literally thousands of subsequent studies.

Our Extension of Attachment Theory into Social Psychology

By the time we began to use attachment theory, in the late 1980s it had been extensively tested in studies of child development, [p. 163 ↓] and Ainsworth's infant attachment categories were well known. For various reasons, including (we believe) the increasing number of women entering social psychology (Elaine Hatfield and Ellen Berscheid being two prominent and highly influential examples), the increasing divorce rate in the US, and a concern with loneliness in industrialized societies (e.g., Peplau and Perlman, 1982), social psychologists were beginning to concern themselves with the formation, maintenance, and dissolution of close relationships. This concern was manifested in the creation of new professional organizations focused on the study of romantic and marital relationships and in a landmark 1983 book, *Close Relationships*, edited by Harold Kelley, one of the most prominent social psychologists of his (or any other) generation, along with Ellen Berscheid and several others. Suddenly the study of

love was not merely professionally acceptable, it was highly visible, even in journals like *Psychological Review* (e.g., Sternberg, 1986).

A problem during that period, at least in our estimation, was the exceptional prominence in social psychology of the attitude construct, which had been central for decades. Its familiarity caused researchers, at first, to consider love to be just another attitude (e.g., Hendrick and Hendrick, 1986). Little consideration was given to the fact that romantic and parental love had existed for millennia (Jankowiak, 1995; Singer, 1987) and that the inherent importance of love and loss could be seen in the lives of nonhuman primates (e.g., Harlow, 1959). These were the days before evolutionary psychology. Moreover, social psychologists were generally unaware that psychoanalysts from Freud to Bowlby had written a great deal about the psychodynamics of filial and romantic love and the relation of romantic love to sexuality. We were unusual among social psychologists in having been deeply interested in psychoanalysis since first encountering it in our undergraduate years.

In our view, anyone who pays close attention to what goes on in people's lives, or who reads romantic novels or poems or studies art or film, realizes that the issues raised by psychoanalysts, beginning with Freud, are crucial: sexual attraction and desire; romantic love and longing; the development of personality in the crucible of family relationships; painful, corrosive emotions such as anger, fear, jealousy, grief, hatred, shame, and remorse, which contribute to intrapsychic conflicts, defenses, and psychopathology; intergroup hostility and war. Given our interests, social psychology at first seemed superficial compared with psychoanalysis. Nevertheless, social psychology's strong point – which was the fatally weak point of psychoanalysis – was the use of experimental methods and creative experimental interventions. Psychoanalytic theorists seemed capable of endlessly inventing hypothetical constructs and invisible mental processes without being constrained by operational definitions, sound psychometrics, or replicable empirical methods. Hence, social psychology was capable of rendering psychodynamic theories testable.

Both of us began our careers as experimental researchers pursuing then-popular topics in social and personality psychology (stress and learned helplessness in Mikulincer's case, self-awareness and fear of success in the case of Shaver), but our interest in psychoanalytic theory never subsided. When we encountered Bowlby's

books, we realized that a psychoanalytic thinker could incorporate the full range of scientific perspectives on human behavior, seek empirical evidence for psychoanalytic propositions, and amend or reformulate psychoanalytic theory based on empirical research. Ainsworth's development of the laboratory Strange Situation assessment procedure, which allowed her to systematically classify infants' attachment patterns and relate them to reliable observations of parent-child interactions at home, added to our confidence that extending attachment theory and its research methods into the realm of adolescent and adult love relationships might be possible.

In the mid 1980s, Shaver was studying adolescent and adult loneliness (see, for example, Rubenstein and Shaver, 1982; Shaver and [p. 164 ↓] Hazan, 1984) and noticing both that attachment theory was useful in conceptualizing loneliness (e.g., Weiss, 1973) and that patterns of chronic loneliness were similar in certain respects to the insecure infant attachment patterns identified by Ainsworth and her colleagues (1978). Building on this insight, one of Shaver's doctoral students, Cindy Hazan, wrote a seminar paper suggesting that attachment theory could be used as a framework for studying romantic love – or “romantic attachment,” as they called it in their initial article on the topic (Hazan and Shaver, 1987).

That article caught the eye of Mikulincer, who had become interested in attachment theory while studying affect-regulation processes related to learned helplessness, depression, combat stress reactions, and post-traumatic stress disorder in Israel. He noticed similarities between (1) certain forms of helplessness in adulthood and the effects of parental unavailability in infancy; (2) intrusive images and emotions in the case of post-traumatic stress disorder and the anxious attachment pattern described by Ainsworth et al. (1978) and Hazan and Shaver (1987); and (3) avoidant strategies for coping with stress and the avoidant attachment pattern described by these same authors. In 1990, Mikulincer, Florian, and Tolmacz published a study of attachment patterns and conscious and unconscious death anxiety, one of the first studies to use the preliminary self-report measure of adult attachment patterns devised by Hazan and Shaver (1987), and the first to show its ability to illuminate unconscious mental processes.

From then on, both of us continued to pursue the application of attachment theory to the study of adults' emotions, emotion-regulation strategies, and close relationships,

noticing that we were both interested in the experimental study of what might be called attachment-related psychodynamics: the kinds of mental processes, including intense needs, powerful emotions and conflicts, and defensive strategies, that had captivated the attention of both Freud and Bowlby. We decided to pool our efforts to craft a more rigorous formulation of adult attachment theory (e.g., Mikulincer and Shaver, 2003, 2007a; Shaver and Mikulincer, 2002), clarify and extend our model of the attachment behavioral system, test the model in many different ways, including the use of priming techniques developed by cognitively oriented social psychologists, and incorporate within our theory some of positive psychology's emphasis on personal growth and social virtues (e.g., Gable and Haidt, 2005; Seligman, 2002) and some of organizational psychology's emphasis on leadership and group dynamics (e.g., Davidovitz et al., 2007; Rom and Mikulincer, 2003). Today, adult attachment theory, as summarized in our 2007 book, is one of the leading approaches to research on social relationships, personality processes, and the psychodynamic nature of the human mind. The current form of the theory is summarized in the following section.

Adult Attachment Theory

What we are calling *adult* attachment theory (Mikulincer and Shaver, 2007a) includes six major constructs: (1) the innate attachment behavioral system (Bowlby, 1982; Mikulincer and Shaver, 2006), (2) attachment-related affect-regulation strategies (Main, 1990; Mikulincer et al., 2008), (3) internal working models of self and others (Bowlby, 1982; Mikulincer and Shaver, 2005), (4) attachment patterns or "styles" (Ainsworth et al., 1978; Shaver and Mikulincer, 2009), (5) attachment security viewed as a resilience resource that enhances self-esteem and promotes prosocial emotions and behavior (Bowlby, 1988; Mikulincer and Shaver, 2007b), and (6) dysfunctional aspects of attachment insecurity (Cassidy and Kobak, 1988; Shaver and Mikulincer, 2002). We will discuss each of these constructs in turn.

Behavioral Systems

To characterize the motives involved in personal and social development, which Freud [p. 165 ↓] had attempted to explain in terms of sexual and aggressive, or life and death, “instincts” or “drives,” Bowlby (1982) took from ethology the concept of *behavioral systems* – species-universal, biologically evolved neural programs that organize behavior in ways that increase the likelihood of an animal’s survival and reproduction. He viewed these systems as similar to cybernetic control systems, which are not powered by drives.

According to Bowlby (1982), one of the earliest behavioral systems to appear in human development is the *attachment system*, whose biological function is to protect a person (especially during infancy and early childhood) from danger by assuring that he or she maintains proximity to caring and supportive others (whom Bowlby, 1982, called *attachment figures*). In Bowlby’s (1982) view, the need to seek out and maintain proximity to attachment figures evolved because of the prolonged dependence of human children on “stronger and wiser” others (often, but not always, the parents), who can defend the children from predators and other dangers while supporting their gradual development of knowledge and skills.

Because human (and other primate) infants seem naturally to look for and gravitate toward *particular* others (those who are familiar and at least sometimes helpful), and to prefer them over alternative caregivers, Bowlby used the terms “affectional bond” and “attachment” for the processes that link one person with another in close relationships. This is the reason he called his theoretical formulation *attachment theory*. Although the attachment system is most important and most evident during the early years of life, Bowlby (1988) claimed that it is active across the lifespan and is most frequently manifested when a person seeks support, affection, or protection from a close relationship partner. This lifespan orientation encouraged developmental and social psychologists (e.g., Main et al., 1985; Shaver et al., 1988) to extend the theory into the domain of adolescent and adult relationships.

During infancy, primary caregivers (usually one or both parents but also grandparents, neighbors, older siblings, daycare workers, and so on) are likely to occupy the role of

attachment figure. Ainsworth (1973) reported that infants tend to seek proximity to their primary caregiver when tired or ill, and Heinicke and Westheimer (1966) found that infants tend to be most easily soothed by their primary caregivers. During adolescence and adulthood, other relationship partners often become targets of proximity seeking and emotional support, including close friends and romantic partners. Teachers and supervisors in academic settings or therapists in clinical settings can also serve as real or potential sources of comfort and support. Moreover, groups, institutions, and symbolic personages (e.g., God, the Buddha, or Virgin Mary) can be used mentally as attachment figures. This array of real and symbolic figures, which can vary in importance or centrality, form what Bowlby (1982) called a *hierarchy of attachment figures*.

According to attachment theory, a particular relationship partner is an attachment figure and a relationship is an attachment relationship only to the extent that the relationship partner accomplishes, or is called upon to accomplish, three important functions (e.g., Ainsworth, 1991; Hazan and Shaver, 1994; Hazan and Zeifman, 1994). First, the attachment figure is someone to whom the attached individual seeks proximity in times of stress or need. Moreover, separation from or loss of this person elicits distress, protest, and efforts to achieve reunion (either literally or, in the case of grief, symbolically). Second, the person is viewed as a real or potential “safe haven,” because he or she provides, or is hoped to provide, comfort, support, protection, and security in times of need. Third, the person is viewed as a “secure base,” allowing a child or adult to pursue nonattachment goals without undue concerns about safety and to sustain exploration, risk taking, and self-development.

What attachment theory calls *activation of the attachment system* can be seen in the [p. 166 ↓] behavior of human infants, who tend to drop whatever they are doing (e.g., playing with interesting toys) and seek comfort and support from an attachment figure if an odd or unexpected noise is heard or a stranger enters the room (Ainsworth et al., 1978). (Bowlby [1982] considered these stimuli, as well as finding oneself in the dark or feeling ill, to be natural clues to danger which have had obvious significance for survival throughout human evolutionary history.) The same kind of attachment-system activation is notable in the minds of adults who are subjected to conscious or unconscious threats. For example, we (Mikulincer et al., 2002) conducted experiments in which we subliminally presented threatening words (e.g., failure, separation) to

adults and then assessed indirectly (using reaction times in a word-identification or Stroop color-naming task) which names of relationship partners became more mentally available for processing following the unconscious threat. It turned out that the names of a person's attachment figures became more available following unconscious exposure to threatening words. These words had no effect on the mental availability of names of people, even familiar ones, who were not viewed as attachment figures. That is, attachment figures are not just any relationship partners; they are special people to whom one turns, even unconsciously and automatically, when comfort or support is needed.

According to Bowlby (1982), the natural goal of the attachment system is to increase a person's sense of *security* (which Sroufe and Waters, 1977, called *felt security* to emphasize its emotional qualities) – a sense that the world is a safe place, that one can rely on others for protection, comfort, and support, and that one can confidently explore one's environment and engage in social and nonsocial activities without undue or debilitating fear of damage. This goal is made particularly salient by encounters with actual or symbolic threats or by appraising an attachment figure as not sufficiently available or responsive. In such cases, the attachment system is activated and the individual is driven to reestablish actual or symbolic proximity to an attachment figure (which attachment researchers call the *primary strategy* of the attachment system; Main, 1990). These bids for proximity persist until the sense of security is restored, at which time the attachment system is deactivated, or downregulated, and the individual can calmly and skillfully return to other activities. That is, the search for protection, support, and security is not only a goal in itself but also an important foundation for attaining non-attachment goals. This makes attachment theory distinct from other social psychological theories that characterize self-esteem, belongingness, cognitive consistency, or social influence or dominance a primary goal. Many of these goals are related to security and insecurity, but we view them as offshoots rather than roots of social motivation.

During infancy, the primary attachment strategy includes tactics such as nonverbal expressions of need (e.g., reaching, crying, pleading) and crawling or toddling toward an attachment figure to increase proximity and safety (Ainsworth et al., 1978). In adulthood, these tactics are expanded to include many other methods of establishing contact (e.g., talking, calling an attachment figure on the telephone, sending an e-

mail or text message) as well as ways of calling upon soothing, comforting mental representations of attachment figures or even self-representations associated with these figures (e.g., qualities of oneself modeled on qualities of an attachment figure or feelings associated with being loved and comforted by such figures; Mikulincer and Shaver, 2004). Many studies (reviewed by Mikulincer and Shaver, 2007b) have demonstrated that “priming” a person with conscious or unconscious reminders or representations of attachment figures can increase felt security and thereby alter a person's feelings about objects, situations, and people, reduce hostility to outgroup members, facilitate empathy, compassion, and altruistic helping, and encourage creative forms of exploration.

[p. 167 ↓]

Bowlby (1988) summarized many of the adaptive benefits of proximity seeking. For example, he viewed successful bids for proximity and the resulting boost in felt security as important to the creation and maintenance of successful, satisfying relationships. Every attachment-related interaction that restores a person's sense of security reaffirms the value of interpersonal closeness and strengthens affectional bonds with the person responsible for boosting one's sense of security. This is how attachment researchers explain many of the effects identified by non-attachment social psychology researchers (e.g., Murray et al., 2006; Reis et al., 2004; Rusbult et al., 1991) as important determinants of relationship stability and quality.

Moreover, successful bids for closeness and emotional support play an important role in helping a person regulate and de-escalate negative emotions such as anger, sadness, anxiety, and demoralization (Bowlby, 1973, 1980). According to attachment theory, successful self-regulation is learned, at first, with the help of attachment figures who accurately perceive one's negative emotions and the situational causes of these emotions and then gently and effectively soothe one's troubled mind and offer useful suggestions for solving problems or reappraising troubling situations. They therefore help a person maintain emotional balance and resilience in the face of stress.

Individual Differences in Attachment-System Functioning

Attachment theory is a general theory of social and emotional development, but it would not have captured the attention of developmental, personality, social, and clinical psychologists if it had been only that. What captured research psychologists' attention were the patterns or styles of attachment emphasized in Bowlby's theory and operationalized in Ainsworth's research on mothers and infants. Most of the research and clinical applications inspired by attachment theory are concerned with those individual differences.

Attachment-Figure Availability and Secondary Attachment Strategies

Besides its species-universal operating characteristics, the attachment behavioral system includes regulatory parameters than can be influenced by a person's history of interactions with attachment figures. In early infancy, the effects of experience can be conceptualized in terms of simple learning principles. If a particular behavioral strategy (e.g., crying for help, protesting angrily, downregulating distress signals) works with a particular caregiver, it will be reinforced. If a particular strategy results in punishment or caregiver withdrawal, it will become less available in an infant's behavioral repertoire (perhaps by being actively suppressed). The same is true for the young of other mammalian species.

In the case of human children, however, what is learned includes not only automatic behavior patterns but also vivid memories, abstracted assumptions, and expectations about caregivers' reactions and the effectiveness or ineffectiveness of one's own possible behaviors. Because Bowlby and Ainsworth were working during what has been called, in retrospect, the cognitive revolution in psychology, they were sensitive to the role played by memories, cognitive schemas, and other mental representations in regulating the attachment system. In attachment theory, these mental structures

and processes are called *internal working models* of self and others (Bowlby, 1982). Over time, a person's working models, which contain both conscious and unconscious elements, become molded by the quality of interactions with attachment figures, thereby "programming" the attachment system to expect and conform to these figures' characteristic behaviors. Through this process, a person learns to adjust his or her attachment system to fit contextual demands and rely on expectations about possible access routes to protection and security. These working models are thought to be the basis of both current individual [p. 168 ↓] differences in attachment strategies, or styles, and within-person continuity in the operation of the attachment system over time.

According to Bowlby (1973, 1988), variations in working models, and hence in attachment-system functioning, depend on the availability, sensitivity, and responsiveness of attachment figures in times of need. When one's key relationship partner is available, sensitive, and responsive to one's proximity- and support-seeking efforts, one is likely to experience "felt security" and thus increase one's confidence in proximity seeking as an effective distress-regulation strategy. During such interactions one also acquires procedural knowledge about distress management, which is organized around what attachment researchers (e.g., Mikulincer et al., 2009) have characterized as a *secure-base script*. This script is thought to include the following if-then propositions: "If I encounter an obstacle and/or become distressed, I can approach a significant other for help; he or she is likely to be available and supportive; I will experience relief and comfort as a result of seeking proximity to this person; and I can then return to other activities."

When a primary attachment figure, however, proves not to be available, sensitive, or responsive, felt security is not attained and the distress that initially activated proximity-seeking efforts is compounded by attachment-related doubts and concerns (e.g., Can I trust others in times of need?). These troubling interactions indicate that the primary attachment strategy (seek proximity and support) is failing to accomplish its goal and that alternative strategies should be adopted to cope with current insecurities and distress. Attachment theorists (e.g., Cassidy and Kobak, 1988; Main, 1990) have called these alternative strategies *secondary attachment strategies*, which (based on Ainsworth et al.'s [1978] research) are thought to take two major organized forms: *hyperactivation* and *deactivation* of the attachment behavioral system.

Hyperactivated strategies include what Bowlby (1982) called *protest* reactions to the frustration of attachment needs. Protest often occurs in relationships in which the attachment figure is sometimes unreliably responsive but sometimes not, placing the needy individual on a partial reinforcement schedule that seems to reward persistence in the use of energetic, strident, noisy proximity-seeking strategies, because such attempts sometimes succeed. In such cases, a person does not give up on proximity-seeking and in fact intensifies it in an effort to demand or coerce the attachment figure's attention, love, and support. The main goal of these strategies is to make an unreliable or insufficiently responsive figure provide support and security. This involves exaggerating threat appraisals and overemphasizing indications of attachment-figure unavailability while intensifying demands for attention, care, and love. This strategy can, paradoxically, be viewed as a form of affect regulation, even though it involves upregulation rather than downregulation – the usually taken-for-granted method of effectively regulating emotions (Gross, 1999).

Deactivating strategies, in contrast, are efforts to escape, avoid, or minimize the pain and frustration caused by unavailable, unsympathetic, or unresponsive attachment figures. This kind of response typically occurs in relationships with attachment figures who disapprove of and punish closeness and expressions of need, dependence, and vulnerability (Ainsworth et al., 1978). In such relationships, a needy individual learns to expect better outcomes if proximity-seeking bids are suppressed, the attachment system is deactivated, and threats and dangers are dealt with on one's own. (Bowlby [1982] called this strategy *compulsive self-reliance*.) The primary goal of deactivating strategies is to keep the attachment system quiescent or downregulated to avoid recurring frustration and distress caused by interactions with cold, neglectful, or punishing attachment figures. Such deactivation requires that a person deny attachment needs, avoid intimacy and interdependence in relationships, and distance him- or herself from threats that might cause [p. 169 ↓] unwanted and potentially unmanageable activation of attachment-related needs, thoughts, feelings, or behaviors.

Internal Working Models

As already mentioned, Bowlby (1982) theorized that memories of important social interactions with attachment figures are stored and eventually schematized in an

associative memory network. This stored knowledge allows a person to predict the likely course and outcomes of future interactions with an attachment figure and to adjust proximity-seeking bids accordingly. Repeated augmentation and editing of these models generally result in increasingly stable mental representations of self, attachment figures, and relationships. Bowlby (1982) wrote about two major forms of working models: representations of attachment figures' responses and inclinations (*working models of others*) and representations of the self's lovability and efficacy (*working models of self*). Once the attachment system has operated for several years in the context of attachment relationships, it is linked with complex representations of the availability, responsiveness, and sensitivity of these figures as well as representations of the self's ability to elicit a partner's attention and affection when desired.

During infancy and childhood, working models are based on specific interactions, or kinds of interactions, with particular attachment figures. As a result, a child can hold multiple episodic (situation- or person-specific) representations of self and others that differ with respect to an interaction's outcome (especially success or failure at gaining felt security), and with respect to the secondary strategy used to deal with insecurity during that interaction (hyperactivating or deactivating). With experience and in the context of cognitive development, these episodic representations form excitatory and inhibitory associations with each other. For example, experiencing or thinking about an episode of security attainment activates memories of similar security-enhancing episodes and renders memories of attachment insecurities and worries less accessible. These associations favor the formation of more abstract and generalized attachment representations with a specific partner. Then, through excitatory and inhibitory links with models representing interactions with other attachment figures, even more generic working models are formed to summarize relationships in general. This process of continual model construction, renovation, and integration results, over time, in the creation of a hierarchical associative network that includes episodic memories, relationship-specific models, and generic working models of self and others. Overall et al. (2003) provided statistical evidence for this hierarchical network of attachment working models.

Unfortunately, the theoretical literature on attachment has sometimes made it seem that working models are simple and univocal with respect to important relationship issues. Research evidence suggests, however, in line with Bowlby's (e.g., 1980) original ideas

about multiple models, conflicting models, and conscious and unconscious models, that most people can remember and be affected by both security-enhancing interactions with attachment figures and security-reducing interactions (e.g., Baldwin et al., 1996; Mikulincer and Shaver, 2007b). It thus matters a great deal what a particular person is reminded of, or is thinking about, when attachment-related processes and outcomes are assessed. The mental representation of one relationship may differ from the mental representation of another, and focusing on a particular issue (e.g., sexual infidelity) may cause related previous experiences to become temporarily more mentally accessible and psychologically influential than usual.

The notion that everyone has multiple attachment models organized by a hierarchical cognitive network raises questions about which model will be most accessible (i.e., readily activated and used to guide attachment-related expectations, defenses, and behaviors) in a given situation. As with other mental representations, the accessibility of an attachment working model is determined [p. 170 ↓] by the amount of experience on which it is based, the number of times it has been applied in the past, the density of its connections with other working models, and the issues made salient in a particular situation (e.g., Shaver et al., 1996). At the relationship-specific level, the model representing the typical interaction with an attachment figure has the highest likelihood of being accessible and guiding subsequent interactions with that person. At the generic level, the model that represents interactions with major attachment figures (e.g., parents and romantic partners) typically becomes the most commonly available representation and has the strongest effect on attachment-related expectations, feelings, and behaviors across relationships and over time.

According to Bowlby (1973), consolidation of a regularly available working model is the most important psychological process accounting for the enduring, long-term effects of attachment interactions during infancy, childhood, and adolescence on attachment-related cognitions and behaviors in adulthood. Given a fairly consistent pattern of interactions with primary caregivers during infancy and childhood, the most representative working models of these interactions become part of a person's implicit procedural knowledge about close relationships, social interactions, and methods of distress regulation; tend to operate automatically and unconsciously; and are resistant to change. Thus, what began as representations of specific interactions with particular primary caregivers during childhood tend to be applied in new situations

and relationships, and eventually have an effect on attachment-related experiences, decisions, and actions even in adulthood (Sroufe et al., 2005).

Beyond the pervasive effects of attachment history on the accessibility of working models, attachment theory also emphasizes, as we have mentioned, the importance of contextual factors that influence the availability of particular models or components of models (e.g., Shaver et al., 1996). Recent studies have shown that contextual cues related to the availability and responsiveness of attachment figures, as well as actual or imagined encounters with supportive or unsupportive figures, can affect which working models become active in memory, even if they are incongruent with a person's more general and more typically available working model (e.g., Mikulincer and Shaver, 2007b; Shaver and Mikulincer, 2008). It seems that the generally accessible and more generic models coexist with less typical working models in a person's associative memory network, and the less typical models can be influenced by contextual factors and are important for understanding a person's behavior in a particular situation (Mikulincer and Shaver, 2007b).

Conceptualization and Measurement of Attachment Patterns or Styles

According to attachment theory (Bowlby, 1988; Shaver and Hazan, 1993), a particular history of attachment experiences and the resulting consolidation of chronically accessible working models lead to the formation of relatively stable individual differences in *attachment style* – the habitual pattern of expectations, needs, emotions, and behavior in interpersonal interactions and close relationships (Hazan and Shaver, 1987). Depending on how it is measured, attachment style characterizes a person's typical attachment-related mental processes and behaviors in a particular relationship (relationship-specific style) or across relationships (global attachment style).

The concept of attachment patterns or styles was first proposed by Ainsworth (1967) to describe infants' patterns of responses to separations from and reunions with their mother at home and in the laboratory Strange Situation procedure, which was designed to activate an infant's attachment system. Based on this procedure, infants were

originally classified into one of three categories: secure, anxious, or avoidant. Main and Solomon (1990) later added a fourth category, *disorganized/disoriented*, characterized by odd, awkward behavior and unusual fluctuations between anxiety and avoidance.

[p. 171 ↓]

Infants classified as secure in the Strange Situation typically react to separation from mother with observable signs of distress, but recover quickly upon reunion with mother and return to exploring the many interesting toys provided in the Strange Situation test room. They greet their mother with joy and affection, initiate contact with her, and respond positively to being picked up and held (Ainsworth et al., 1978). Avoidant infants' reactions are dramatically different and seem to indicate attachment-system deactivation. These infants express little distress when separated from mother and may actively turn away from or avoid her upon reunion. Anxious infants' reactions are hyperactive; these infants cry and protest angrily during separation and show angry, resistant, hyperaroused reactions (i.e., Bowlby's "protest") upon reunion, making it difficult for them to calm down and return to creative play.

In the 1980s, researchers from different psychological subdisciplines (developmental, clinical, personality, and social) constructed new attachment measures to extend attachment theory into adolescence and adulthood. Based on a developmental and clinical approach, Main and her colleagues (George et al., 1985; Main et al., 1985; see Hesse, 2008, for a review) devised the Adult Attachment Interview (AAI) to study adolescents and adults' mental representations of attachment to their parents during childhood. One of the major findings of this approach to studying adult attachment is that an adult's AAI classification (secure, dismissing, preoccupied, or unresolved) predicts his or her infant child's attachment pattern in the Strange Situation (see van IJzendoorn, 1995, for a review), even if the interview is conducted before the infant was born. In other words, there is good evidence for the intergenerational transmission of attachment patterns, which seems not to be primarily a matter of shared genes (e.g., Vaughn et al., 2008).

In an independent line of research, Hazan and Shaver (1987), who wished to apply Bowlby and Ainsworth's ideas to the study of romantic relationships, developed a self-report measure of adult attachment style. In its original form, the measure consisted of

three brief descriptions of constellations of feelings and behaviors in close relationships that were intended to parallel the three infant attachment patterns identified by Ainsworth et al. (1978). College students and community adults were asked to read the three descriptions and place themselves in one of the three attachment categories according to their predominant feelings and behavior in romantic relationships. The three descriptions were as follows:

- 1. *Secure*: I find it relatively easy to get close to others and am comfortable depending on them and having them depend on me. I don't worry about being abandoned or about someone getting too close to me.
- 2. *Avoidant*: I am somewhat uncomfortable being close to others; I find it difficult to trust them completely, difficult to allow myself to depend on them. I am nervous when anyone gets too close and often, others want me to be more intimate than I feel comfortable being.
- 3. *Anxious*: I find that others are reluctant to get as close as I would like. I often worry that my partner doesn't really love me or won't want to stay with me. I want to get very close to my partner and this sometimes scares people away.

Hazan and Shaver's (1987) study was followed by scores of others that used the simple forced-choice self-report measure to examine the interpersonal and intrapersonal correlates of adult attachment style (see reviews by Shaver and Hazan, 1993; Shaver and Mikulincer, 2002). Over time, attachment researchers made methodological and conceptual improvements to the original self-report measure and reached the conclusion that attachment styles are best conceptualized as regions in a two-dimensional space (e.g., Bartholomew and Horowitz, 1991; Brennan et al., 1998). The first dimension, which we call attachment-related *avoidance*, is concerned with discomfort with closeness and with dependence on relationship partners and a preference for emotional distance and self-reliance. Avoidant individuals identified [p. 172 ↓] with self-report measures use deactivating attachment and affect-regulation strategies to deal with insecurity and distress. The second dimension, attachment-related *anxiety*, includes a strong desire for closeness and protection, intense worries about one's partner's availability and responsiveness and one's own value to the partner, and the use of hyperactivating strategies for dealing with insecurity and

distress. People who score low on both dimensions are said to be secure or to have a secure attachment style.

The two attachment-style dimensions can be measured with the 36-item Experiences in Close Relationships scale (ECR; Brennan et al., 1998), which is reliable in both the internal-consistency and test-retest senses and has high construct, predictive, and discriminant validity (Crowell et al., 2008). Eighteen items tap the avoidance dimension (e.g., "I try to avoid getting too close to my partner," "I prefer not to show a partner how I feel deep down"), and the remaining 18 items tap the anxiety dimension (e.g., "I need a lot of reassurance that I am loved by my partner," "I resent it when my partner spends time away from me"). The two scales were conceptualized as independent and have been found to be empirically uncorrelated or only weakly correlated in most studies. Studies based on self-report measures of adult attachment style, some based on three categories, some on four categories (including two kinds of avoidance, labeled fearful and dismissive), and some on two dimensions, have allowed researchers to document theoretically predictable attachment-style variations in relationship quality, mental health, social adjustment, ways of coping, emotion regulation, self-esteem, interpersonal behavior, and social cognitions (see Mikulincer and Shaver, 2003, 2007a, for reviews).

There is relatively little research on the heritability of the two major forms of attachment insecurity measured with self-report scales (e.g., Crawford et al., 2007; Donnellan et al., 2008) or with the Adult Attachment Interview (Torgersen et al., 2007). The self-report studies provide fairly consistent evidence of genetic influences on attachment anxiety, which is correlated with neuroticism (e.g., Nofle and Shaver, 2006), a personality trait influenced by genes. The evidence for genetic influences on avoidant attachment is less consistent, but Gillath et al. (2008) found that both attachment anxiety and avoidance were associated with particular genetic alleles: anxiety with a polymorphism of the DRD2 dopamine receptor gene and avoidance with a polymorphism of the 5HT2A serotonin receptor gene. This line of research is still in its infancy, but none of the studies published to date suggest that genes are likely to be the most important determinants of attachment style.

Attachment-Figure Availability and the Broaden-and-Build Cycle of Attachment Security

Having outlined attachment theory's major constructs and some of the methods and procedures used to operationalize them, we turn to some of the personal, dyadic, and social-system consequences of variations in attachment-system functioning. In the present section of the chapter we are especially interested in beneficial effects of perceived attachment-figure availability and the resulting sense of security on social judgments, self-image, personality development, mental health, and relationship quality. After considering the beneficial effects of security, we will turn to defensive processes related to secondary attachment strategies (anxious hyperactivation and avoidant deactivation) and to the emotional and adjustment difficulties that arise when a person relies on defensive processes.

According to attachment theory, the physical and emotional availability of an actual security provider, or ready access to mental representations of supportive attachment figures, results in a sense of felt security and fosters what we, following Fredrickson (2001), call a *broaden-and-build cycle* of [p. 173 ↓] attachment security. This cycle is a cascade of mental and behavioral processes that can be viewed as resources for maintaining emotional stability in times of stress, fostering open and deeply interdependent bonds with others, optimizing personal adjustment, and expanding one's perspectives and capacities. In the long run, repeated experiences of attachment-figure availability have enduring effects on intrapsychic organization and interpersonal behavior. At the intrapsychic level, such experiences can be called upon as resilience resources, sustaining emotional wellbeing and personal adjustment, and resulting in positive working models of self and others that are readily accessible in memory when needed to bolster a person's mood and coping capacity. At the interpersonal level, repeated experiences of attachment-figure availability allow a person to develop the skills and attitudes associated with a secure attachment style, which facilitates the formation and maintenance of warm, satisfying, stable, and harmonious relationships.

The most immediate psychological effect of having reliable, dependable access to an available, sensitive, and responsive attachment figure is effective management of distress and relatively rapid restoration of emotional equanimity following threats and stresses. As a result of good relationships with attachment figures, secure people remain relatively unperturbed in times of stress and experience longer periods of positive affectivity, which contribute to stable mental health. Indeed, several studies have shown that secure attachment is associated positively with measures of wellbeing and negatively with measures of hurt feelings, negative affectivity, depression, and anxiety (for reviews, see Mikulincer and Shaver, 2003, 2007a; Shaver and Mikulincer, 2002, 2009).

Experiences of attachment-figure availability also contribute to the construction of an extensive network of positive mental representations, which plays an important role in maintaining emotional stability and adjustment. One part of this network concerns the appraisal of life's problems as manageable, which helps a person maintain an optimistic and hopeful stance. Relatively secure people can appraise and reappraise stressful events in positive ways and thereby deal more effectively with them. Researchers have consistently found positive correlations between self-reports of attachment security and constructive, optimistic appraisals of stressful events (see Mikulincer and Shaver, 2007a, for an extensive review).

Another set of security-related mental representations concerns other people's intentions and traits. Numerous studies have shown that more securely attached people hold more positive views of human nature, use more positive terms when describing relationship partners, perceive relationship partners as more supportive, have more positive expectations concerning their partners' behavior and tend to explain a partner's hurtful behavior in less negative ways (for reviews, see Mikulincer and Shaver, 2003; Shaver and Mikulincer, 2006).

Interactions with available and sensitive relationship partners reduce worries about being rejected, criticized, or abused. Such interactions confirm that a caring partner is unlikely to betray one's trust, react coldly or abusively to one's expressions of need, or respond unfavorably to bids for closeness and comfort. Numerous studies confirm that secure individuals score higher on measures of self-disclosure, support

seeking, intimacy, trust, open communication, prorelational behavior, and relationship satisfaction (for reviews, see Feeney, 2008; and Shaver and Mikulincer, 2006).

Interactions with security-enhancing attachment figures also strengthen a person's authentically positive sense of self-esteem and social value (Mikulincer and Shaver, 2003). That is, secure individuals generally feel safe and protected and perceive themselves as valuable, lovable, and special, thanks to being valued, loved, and regarded as special by caring relationship partners. Research consistently shows that secure individuals have higher self-esteem (e.g., Bartholomew and Horowitz, 1991) and view themselves as [p. 174 ↓] more competent and efficacious (e.g., Cooper et al., 1998) than insecure individuals.

A relatively secure person's resources for dealing with stress make it less necessary for them to rely on psychological defenses that distort perception, limit coping flexibility, and generate interpersonal conflict. A secure person can devote mental resources to personal growth that would otherwise have to be devoted to preventive, defensive maneuvers; they can also attend to other people's needs and feelings rather than, or in addition to, their own. Being confident that support is available if needed, a secure person can take calculated risks and accept important challenges that contribute to the broadening of perspectives and skills, which is an important part of personal growth. Indeed, research indicates that attachment security is associated with enhanced curiosity and learning, encourages relaxed exploration of new, unusual information and phenomena, and favors the formation of open and flexible cognitive structures, despite the uncertainty and confusion that a broadening of experience might entail (e.g., Elliot and Reis, 2003).

Attachment security is associated with higher scores on self-report measures of responsiveness to a relationship partner's needs (e.g., Kuncz and Shaver, 1994) and with more supportive reactions to a distressed relationship partner (e.g., Simpson et al., 1992). Both dispositional and contextually augmented attachment security have also been associated with heightened compassion for a suffering individual and willingness to relieve the person's distress (e.g., Gillath et al., 2005; Mikulincer et al., 2005).

Secondary Attachment Strategies, Emotional Difficulties, and Psychological Maladjustment

According to attachment theory (Main, 1990; Mikulincer and Shaver, 2003, 2007a; Shaver and Mikulincer, 2002), secondary attachment strategies (anxious hyperactivation and avoidant deactivation) are defenses against the frustration and pain caused by the unavailability, unreliability, or unresponsiveness of attachment figures in times of need. Although these secondary strategies are initially aimed at achieving a workable relationship with an inconsistently available or consistently distant or unavailable attachment figure, they are maladaptive when used in later relationship situations in which proximity-seeking, psychological intimacy, and collaborative interdependence would be more productive and rewarding. Moreover, these strategies result in the maintenance of distorted or constraining working models and affect-regulation techniques that are likely to interfere with psychological health, personal growth, and social adjustment.

Anxious attachment encourages distress intensification and the arousal of negative memories, expectations, and emotions, which in turn interfere with mental coherence and, in some cases, precipitate episodes of serious psychopathology (Mikulincer and Shaver, 2003). Although avoidant people can maintain a defensive façade of security and imperturbability, they ignore, misinterpret, or misunderstand their own emotions and have difficulty dealing with prolonged, demanding stressors that require active problem confrontation and mobilization of external sources of support (Mikulincer and Shaver, 2003). In addition, although avoidant people are able to consciously suppress or ignore distress, the distress can still be indirectly expressed in somatic symptoms, sleep disturbances, and reduced immunity to diseases. Moreover, avoidant individuals can transform personal distress into feelings of hostility, loneliness, and estrangement from others (Shaver and Hazan, 1993).

Many studies confirm that attachment anxiety is inversely related to wellbeing and positively associated with global distress, depression, anxiety, eating disorders,

substance abuse, conduct disorders, and severe personality disorders (see Mikulincer and Shaver, 2007a, for a review). With regard to [p. 175 ↓] avoidant attachment, many studies have found no significant associations between avoidant attachment and *self-report measures* of wellbeing and global distress (see Mikulincer and Shaver, 2007a, for a review). However, several studies indicate that avoidant attachment is associated with particular patterns of emotional and behavioral problems, such as a pattern of depression characterized by perfectionism, self-punishment, and self-criticism, somatic complaints; substance abuse and conduct disorders; and schizoid and avoidant personality disorders (see Mikulincer and Shaver, 2007a, for a review). In addition, whereas no consistent association has been found in community samples between avoidant attachment and global distress, studies that focus on highly demanding and stressful events, such as giving birth to a seriously handicapped infant, reveal that avoidance is related to higher levels of distress and poorer long-term outcomes (e.g., Berant et al., 2008).

Attachment Theory: Relation to other Theories and Potential Applications

Attachment theory interfaces with and has much in common with other theories discussed in this handbook. Attachment theory is, for example, an early evolutionary-psychological theory, having been constructed partly with reference to ethological studies of birds and nonhuman primates. It is closely related to interdependence theories of close relationships, which focus on interpersonal transactions and emphasize the influence of one person's responses on another's person's outcomes. Both attachment and interdependence theories emphasize the importance of trust between relationship partners. Attachment theory is not, however, exclusively relational. It includes the important idea that interactions with attachment figures can be biased by defensive processes related to secondary attachment strategies. Because of such biases, working models of the self and others do not exclusively reflect the ways in which a person and relationship partners actually behave in a particular interaction. Rather, they are reflections of both actual social encounters and subjective biases resulting from already well-established defensive strategies. Moreover, attachment-

system activation in adulthood can occur in the mind without necessarily being expressed directly in behavior, and without necessarily requiring the presence of an actual relationship partner.

Both attachment theory and social-cognition theories (included those discussed in the Cognitive Level of Analysis section of this volume) emphasize the extent to which people subjectively construe person–environment transactions, store representations of typical transactions, and use these representations to understand new transactions and organize action plans. In both theoretical approaches, these mental representations guide and coordinate emotion regulation, self-images, person perception, and cognitions, goals, feelings, and behavior in interpersonal settings. Furthermore, attachment theory conceptualizes working models in the same way that social-cognition theorists conceptualize mental representations: they are stored in an associative memory network, maintain excitatory and inhibitory connections with other representations, have a particular level of accessibility determined by past experiences and other factors, and this accessibility can be heightened in a given situation by relevant contextual cues.

Despite these commonalities, however, it would be a mistake to equate attachment working models with the cognitive structures usually studied in social cognition research. In their review of the nature, content, structure, and functions of attachment working models, Shaver et al. (1996) enumerated four differences between these constructs. As compared to other mental representations, (1) working models also tend to deal with a person's wishes, fears, conflicts, and psychological defenses, and they can be affected by these psychodynamic processes; (2) working [p. 176 ↓] models seem to have a larger and more powerful affective component than most social schemas and tend to be shaped more by emotion-regulation processes; (3) working models tend to be construed in more relational terms and to organize representations of the self, others, and social interactions in a highly interdependent fashion; and (4) working models are broader, richer, and more complex structures, and can include tandem or opposite representations of the same person–environment transaction at episodic, semantic, and procedural levels of encoding.

Throughout this chapter we have referred implicitly and explicitly to a wealth of applications of attachment theory and research. In our own research, for example,

we have shown that security enhancement, whether accomplished consciously or unconsciously, has a number of laudable prosocial effects: reducing dogmatism, intolerance of ambiguity, and intergroup hostility; increasing empathy, compassion, and altruism; and increasing participation in community activities. Like self-affirmation procedures (Sherman and Cohen, 2006), which are thought to enhance self-integrity and warm feelings toward other people (Crocker et al., 2008), security-priming procedures reduce perceived threats to the self and make it easier to appreciate and assist others. In several studies we have shown the relevance of attachment theory and research to relationship and marital satisfaction, leadership development, group dynamics, and organizational functioning (Mikulincer and Shaver, 2007a). Attachment research from the beginning was intended to be applicable to parenting education and individual psychotherapy. It is now being applied in marital therapy as well.

Stepping back from the details of the theory and the thousands of empirical studies it has inspired, it is clear that every level of social life would be enhanced, and would be less destructive, violent, and depressing, if people were raised by responsive parents, taught by responsive teachers and mentors, living with responsive spouses, and supervised and guided by responsive leaders. Many new interventions are being created based on attachment research, and so far their track record is very encouraging. In our opinion, attachment theory and research are helping to fulfill the original goal of social psychology, to provide a scientific basis for improving individual and social life.

References

- Ainsworth, M.D.S. (1940) An evaluation of adjustment based on the concept of security . Unpublished doctoral dissertation, University of Toronto, Ontario, Canada.
- Ainsworth, M.D.S. (1967) Infancy in Uganda: Infant Care and the Growth of Love . Baltimore, MD: Johns Hopkins University Press.
- Ainsworth, M.D.S. (1973) The development of infant-mother attachment . In B.M. Caldwell, ed. and H.N. Ricciuti (eds), Review of Child Development Research , 3, 1–94. Chicago: University of Chicago Press.

Ainsworth, M.D.S. (1991) Attachment and other affective bonds across the life cycle . In C.M. Parkes, ed. , J. Stevenson-Hinde, ed. , and P. Marris (eds), *Attachment Across the Life Cycle* , pp. 33–51. New York: Routledge.

Ainsworth, M.D.S., Blehar, M.C., Waters, E. and Wall, S. (1978) Patterns of Attachment: Assessed in the Strange Situation and at Home . Hillsdale, NJ: Erlbaum.

Baldwin, M.W., Keelan, J.P.R., Fehr, B., Enns, V. and Koh Rangarajoo, E. Social-cognitive conceptualization of attachment working models: Availability and accessibility effects . *Journal of Personality and Social Psychology* , 71, 94–109. (1996) <http://dx.doi.org/10.1037/0022-3514.71.1.94>

Bartholomew, K. and Horowitz, L.M. Attachment styles among young adults: A test of a four-category model . *Journal of Personality and Social Psychology* , 61, 226–244. (1991) <http://dx.doi.org/10.1037/0022-3514.61.2.226>

Berant, E., Mikulincer, M. and Shaver, P.R. Mothers' attachment style, their mental health, and their children's emotional vulnerabilities: A seven-year study of children with congenital heart disease . *Journal of Personality* , 76, 31–66. (2008) <http://dx.doi.org/10.1111/j.1467-6494.2007.00479.x>

Bowlby, J. The nature of the child's tie to his mother . *International Journal of Psychoanalysis* , 39, 350–373. (1958)

Bowlby, J. (1973) *Attachment and Loss: Vol. 2. Separation: Anxiety and Anger* . New York: Basic Books.

Bowlby, J. (1979) *The Making and Breaking of Affectional Bonds* . London: Tavistock.

Bowlby, J. (1980) *Attachment and Loss: Vol. 3. Sadness and Depression* . New York: Basic Books.

Bowlby, J. (1982) *Attachment and Loss: Vol. 1. Attachment* , 2nd Edition. New York: Basic Books. (Originally published 1969).

Bowlby, J. (1988) *A Secure Base: Clinical Applications of Attachment Theory*. London: Routledge.

Brennan, K.A., Clark, C.L. and Shaver, P.R. (1998) Self-report measurement of adult romantic attachment: An integrative overview. In J.A. Simpson, ed. and W.S. Rholes (eds), *Attachment Theory and Close Relationships*, pp. 46–76. New York: Guilford Press.

Cassidy, J. and Kobak, R.R. (1988) Avoidance and its relationship with other defensive processes. In J. Belsky, ed. and T. Nezworski (eds), *Clinical Implications of Attachment*, pp. 300–323. Hillsdale, NJ: Erlbaum.

Cooper, M.L., Shaver, P.R. and Collins, N.L. Attachment styles, emotion regulation, and adjustment in adolescence. *Journal of Personality and Social Psychology*, 74, 1380–1397. (1998) <http://dx.doi.org/10.1037/0022-3514.74.5.1380>

Crawford, T.N., Livesley, W.J., Jang, K.L., Shaver, P.R., Cohen, P. and Ganiban, J. Insecure attachment and personality disorder: A twin study of adults. *European Journal of Personality*, 21, 191–208. (2007) <http://dx.doi.org/10.1002/per.602>

Crocker, J., Niiya, Y. and Mischkowski, D. Why does writing about important values reduce defensiveness? Self-affirmation and the role of positive, other-directed feelings. *Psychological Science*, 19, 740–747. (2008) <http://dx.doi.org/10.1111/j.1467-9280.2008.02150.x>

Crowell, J.A., Fraley, R.C. and Shaver, P.R. (2008) Measurement of adult attachment. In J. Cassidy, ed. and P.R. Shaver (eds), *Handbook of Attachment: Theory, Research, and Clinical Applications*, 2nd Edition, pp. 599–634. New York: Guilford Press.

Davidovitz, R., Mikulincer, M., Shaver, P.R., Izsak, R. and Popper, M. Leaders as attachment figures: Leaders' attachment orientations predict leadership-related mental representations and followers' performance and mental health. *Journal of Personality and Social Psychology*, 93, 632–650. (2007) <http://dx.doi.org/10.1037/0022-3514.93.4.632>

de Wolff, M. and van IJzendoorn, M.H. Sensitivity and attachment: A meta-analysis on parental antecedents of infant attachment . *Child Development* , 68, 571–591. (1997) <http://dx.doi.org/10.2307/1132107>

Donnellan, M.B., Burt, S.A., Levendosky, A.A. and Klump, K.L. Genes, personality, and attachment in adults: A multivariate behavioral genetic analysis . *Personality and Social Psychology Bulletin* , 34, 3–16. (2008) <http://dx.doi.org/10.1177/0146167207309199>

Elliot, A.J. and Reis, H.T. Attachment and exploration in adulthood . *Journal of Personality and Social Psychology* , 85, 317–331. (2003) <http://dx.doi.org/10.1037/0022-3514.85.2.317>

Feeley, J.A. (2008). Adult romantic attachment and couple relationships . In J. Cassidy, ed. and P.R. Shaver (eds), *Handbook of Attachment: Theory, Research, and Clinical Applications* , 2nd Edition pp. 456–481. New York: Guilford Press.

Fredrickson, B.L. The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions . *American Psychologist* , 56, 218–226. (2001) <http://dx.doi.org/10.1037/0003-066X.56.3.218>

Gable, S.L. and Haidt, J. What (and why) is positive psychology? *Review of General Psychology* , 9, 103–110. (2005) <http://dx.doi.org/10.1037/1089-2680.9.2.103>

George, C., Kaplan, N. and Main, M. (1985) The Adult Attachment Interview . Unpublished protocol, Department of Psychology, University of California, Berkeley.

Gillath, O., Shaver, P.R., Baek, J. and Chun, D.S. Genetic correlates of adult attachment style . *Personality and Social Psychology Bulletin* , 34, 1396–1405. (2008) <http://dx.doi.org/10.1177/0146167208321484>

Gillath, O., Shaver, P.R., Mikulincer, M., Nitzberg, R.A., Erez, A. and van IJzendoorn, M.H. Attachment, caregiving, and volunteering: Placing volunteerism in an attachment-theoretical framework . *Personal Relationships* , 12, 425–446. (2005) <http://dx.doi.org/10.1111/j.1475-6811.2005.00124.x>

Gross, J.J. (1999) Emotion and emotion regulation . In O.P. John, ed. and L.A. Pervin (eds), *Handbook of Personality: Theory and Research* , 2nd Edition, pp. 525–552. New York: Guilford Press.

Harlow, H.F. Love in infant monkeys . *Scientific American* , 200, 68–86. (1959) <http://dx.doi.org/10.1038/scientificamerican0659-68>

Hazan, C. and Shaver, P.R. Romantic love conceptualized as an attachment process . *Journal of Personality and Social Psychology* , 52, 511–524. (1987) <http://dx.doi.org/10.1037/0022-3514.52.3.511>

Hazan, C. and Shaver, P.R. Attachment as an organizational framework for research on close relationships . *Psychological Inquiry* , 5, 1–22. (1994) http://dx.doi.org/10.1207/s15327965pli0501_1

Hazan, C. and Zeifman, D. (1994) Sex and the psychological tether . In K. Bartholomew, ed. and D. Perlman (eds), *Advances in Personal Relationships: Attachment Processes in Adulthood* , 5, 151–177. London: Jessica Kingsley.

Heinicke, C. and Westheimer, I. (1966) Brief Separations . New York: International Universities Press.

Hendrick, C. and Hendrick, S.S. A theory and method of love . *Journal of Personality and Social Psychology* , 50, 392–402. (1986) <http://dx.doi.org/10.1037/0022-3514.50.2.392>

Hesse, E. (2008) The Adult Attachment Interview: Protocol, method of analysis, and empirical studies . In J. Cassidy, ed. and P.R. Shaver (eds), *Handbook of Attachment: Theory, Research, and Clinical Applications* , 2nd Edition pp. 552–598. New York: Guilford Press.

Hinde, R. (1966) Animal Behavior: A Synthesis of Ethology and Comparative Psychology . New York: McGraw-Hill.

Jankowiak, W.R. (1995) Romantic Passion . New York: Columbia University Press.

KelleyH.H., Berscheid, E., Christensen, A., Harvey, J.H., Huston, T.L., Levinger, G., McClintock, E., Peplau, L.A. and Peterson, D.R. (1983) *Close Relationships* . New York: W.H. Freeman.

Kunce, L.J. and Shaver, P.R. (1994) An attachment-theoretical approach to caregiving in romantic relationships . In K. Bartholomew, ed. and D. Perlman (eds), *Advances in Personal Relationships* , 5, 205–237. London: Kingsley.

Lorenz, K.Z. (1952) *King Solomon's Ring* . New York: Crowell.

Main, M. Cross-cultural studies of attachment organization: Recent studies, changing methodologies, and the concept of conditional strategies . *Human Development* , 33, 48–61. (1990) <http://dx.doi.org/10.1159/000276502>

Main, M., Kaplan, N. and Cassidy, J. Security in infancy, childhood, and adulthood: A move to the level of representation . *Monographs of the Society for Research in Child Development* , 50, 66–104. (1985) <http://dx.doi.org/10.2307/3333827>

Main, M. and Solomon, J. (1990) Procedures for identifying infants as disorganized/disoriented during the Ainsworth strange situation . In M.T. Greenberg, ed. , D. Cicchetti, ed. , and M. Cummings (eds), *Attachment in the Preschool Years: Theory , Research , and Intervention* , pp. 121–160. Chicago: University of Chicago Press.

Mikulincer, M., Florian, V. and Tolmacz, R. Attachment styles and fear of personal death: A case study of affect regulation . *Journal of Personality and Social Psychology* , 58, 273–280. (1990) <http://dx.doi.org/10.1037/0022-3514.58.2.273>

Mikulincer, M., Gillath, O. and Shaver, P.R. Activation of the attachment system in adulthood: Threat-related primes increase the accessibility of mental representations of attachment figures . *Journal of Personality and Social Psychology* , 83, 881–895. (2002) <http://dx.doi.org/10.1037/0022-3514.83.4.881>

Mikulincer, M. and Shaver, P.R. (2003) The attachment behavioral system in adulthood: Activation, psychodynamics, and interpersonal processes . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 35, 53–152. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2803%2901002-5>

Mikulincer, M. and Shaver, P.R. (2004) Security-based self-representations in adulthood: Contents and processes . In W.S. Rholes, ed. and J.A. Simpson (eds), Adult Attachment: Theory, Research, and Clinical Implications , pp. 159–195. New York: Guilford Press.

Mikulincer, M. and Shaver, P.R. (2005) Mental representations of attachment security: Theoretical foundation for a positive social psychology . In M.W. Baldwin (ed.), Interpersonal Cognition , pp. 233–266. New York: Guilford Press.

Mikulincer, M. and Shaver, P.R. (2006) The behavioral system construct: A useful tool for building an integrative model of the social mind . In P.A.M. Van Lange (ed.), Bridging Social Psychology: Benefits of Transdisciplinary Approaches , pp. 279–284. Mahwah, NJ: Erlbaum.

Mikulincer, M. and Shaver, P.R. (2007a) Attachment in Adulthood: Structure, Dynamics, and Change . New York: Guilford Press.

Mikulincer, M. and Shaver, P.R. Boosting attachment security to promote mental health, prosocial values, and inter-group tolerance . Psychological Inquiry , 18, 139–156.
(2007b) <http://dx.doi.org/10.1080/10478400701512646>

Mikulincer, M., Shaver, P.R., Cassidy, J. and Berant, E. (2008) Attachment-related defensive processes . In J.H. Obegi, ed. and E. Berant (eds), Clinical Applications of Adult Attachment , pp. 293–327. New York: Guilford Press.

Mikulincer, M., Shaver, P.R., Gillath, O. and Nitzberg, R.A. Attachment, caregiving, and altruism: Boosting attachment security increases compassion and helping . Journal of Personality and Social Psychology , 89, 817–839. (2005) <http://dx.doi.org/10.1037/0022-3514.89.5.817>

Mikulincer, M., Shaver, P.R., Sapir-Lavid, Y. and Avihou-Kanza, N. What's inside the minds of securely and insecurely attached people? The secure-base script and its associations with attachment-style dimensions . Journal of Personality and Social Psychology , 97, 615–633. (2009) <http://dx.doi.org/10.1037/a0015649>

Murray, S.L., Holmes, J.G. and Collins, N.L. The relational signature of felt security . Psychological Bulletin , 132, 641–666. (2006) <http://dx.doi.org/10.1037/0033-2909.132.5.641>

Noftle, E.E. and Shaver, P.R. Attachment dimensions and the Big Five personality traits: Associations and comparative ability to predict relationship quality . Journal of Research in Personality , 40, 179–208. (2006) <http://dx.doi.org/10.1016/j.jrp.2004.11.003>

Overall, N.C., Fletcher, G.J.O. and Friesen, M.D. Mapping the intimate relationship mind: Comparisons between three models of attachment representations . Personality and Social Psychology Bulletin , 29, 1479–1493. (2003) <http://dx.doi.org/10.1177/0146167203251519>

Peplau, L.A., ed. and Perlman, D. (eds) (1982) Loneliness: A Sourcebook of Current Theory, Research, and Therapy . New York: Wiley.

Reis, H.T., Clark, M.S. and Holmes, J.G. (2004) Perceived partner responsiveness as an organizing construct in the study of intimacy and closeness . In D.J. Mashek, ed. and A. Aron (eds), Handbook of Closeness and Intimacy , pp. 201–225. Mahwah, NJ: Erlbaum.

Rom, E. and Mikulincer, M. Attachment theory and group processes: The association between attachment style and group-related representations, goals, memories, and functioning . Journal of Personality and Social Psychology , 84, 1220–1235. (2003) <http://dx.doi.org/10.1037/0022-3514.84.6.1220>

Rubenstein, C. and Shaver, P.R. (1982) The experience of loneliness . In L.A. Peplau, ed. and D. Perlman (eds), Loneliness: A Sourcebook of Current Theory, Research, and Therapy , pp. 206–223. New York: Wiley.

Rusbult, C.E., Verette, J., Whitney, G.A., Slovik, L.F. and Lipkus, I. Accommodation processes in close relationships: Theory and preliminary empirical evidence . Journal of Personality and Social Psychology , 60, 53–78. (1991) <http://dx.doi.org/10.1037/0022-3514.60.1.53>

Seligman, M.E.P. (2002) *Authentic Happiness: Using the New Positive Psychology to Realize Your Potential for Lasting Fulfillment*. New York: Free Press.

Shaver, P.R., Collins, N.L. and Clark, C.L. (1996) Attachment styles and internal working models of self and relationship partners . In G.J.O. Fletcher, ed. and J. Fitness (eds), *Knowledge Structures in Close Relationships: A Social Psychological Approach* , pp. 25–61. Mahwah, NJ: Erlbaum.

Shaver, P.R. and Hazan, C. (1984) Incompatibility, loneliness, and limerence . In W. Ickes (ed.), *Compatible and Incompatible Relationships* , pp. 163–184. New York: Springer-Verlag.

Shaver, P.R. and Hazan, C. (1993) Adult romantic attachment: Theory and evidence . In D. Perlman, ed. and W. Jones (eds), *Advances in Personal Relationships* , 4, pp. 29–70. London: Jessica Kingsley.

Shaver, P.R., Hazan, C. and Bradshaw, D. (1988) Love as attachment: The integration of three behavioral systems . In R.J. Sternberg, ed. and M. Barnes (eds), *The Psychology of Love* , pp. 68–99. New Haven: Yale University Press.

Shaver, P.R. and Mikulincer, M. Attachment-related psychodynamics . *Attachment and Human Development* , 4, 133–161. (2002) <http://dx.doi.org/10.1080/14616730210154171>

Shaver, P.R. and Mikulincer, M. (2006) Attachment theory, individual psychodynamics, and relationship functioning . In D. Perlman, ed. and A. Vangelisti (eds), *Handbook of Personal Relationships* , pp. 251–272. New York: Cambridge University Press.

Shaver, P.R. and Mikulincer, M. (2008) Augmenting the sense of security in romantic, leader-follower, therapeutic, and group relationships: A relational model of psychological change . In J.P. Forgas, ed. and J. Fitness (eds), *Social Relationships: Cognitive, Affective, and Motivational Processes* , pp. 55–74. New York: Psychology Press.

Shaver, P.R. and Mikulincer, M. (2009) Attachment theory and attachment styles . In M.R. Leary, ed. and R.H. Hoyle (eds), *Handbook of Individual Differences in Social Behavior* , pp. 62–81. New York: Guilford Press.

Sherman, D.K. and Cohen, G.L. (2006) The psychology of self-defense: Self-affirmation theory . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 38, 183–242. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2806%2938004-5>

Simpson, J.A., Rholes, W.S. and Nelligan, J.S. Support seeking and support giving within couples in an anxiety-provoking situation: The role of attachment styles . *Journal of Personality and Social Psychology* , 62, 434–446. (1992) <http://dx.doi.org/10.1037/0022-3514.62.3.434>

Singer, I. (1987) *The Nature of Love: Courtly and Romantic* . Chicago: University of Chicago Press.

Sroufe, L.A., Egeland, B., Carlson, E. and Collins, W.A. (2005) *The Development of the Person: The Minnesota Study of Risk and Adaptation from Birth to Adulthood* . New York: Guilford Press.

Sroufe, L.A. and Waters, E. Attachment as an organizational construct . *Child Development* , 48, 1184–1199. (1977) <http://dx.doi.org/10.2307/1128475>

Sternberg, R.J. A triangular theory of love . *Psychological Review* , 93, 119–135. (1986) <http://dx.doi.org/10.1037/0033-295X.93.2.119>

Torgersen, A.M., Grova, B.K. and Sommerstad, R. A pilot study of attachment patterns in adult twins . *Attachment and Human Development* , 9, 127–138. (2007) <http://dx.doi.org/10.1080/14616730701349705>

van IJzendoorn, M. Adult attachment representations, parental responsiveness, and infant attachment: A meta-analysis on the predictive validity of the Adult Attachment Interview . *Psychological Bulletin* , 117, 387–403. (1995) <http://dx.doi.org/10.1037/0033-2909.117.3.387>

Vaughn, B.E., Bost, K.K. and van IJzendoorn, M.H. (2008) Attachment and temperament: Additive and interactive influences on behavior, affect, and cognition during infancy and childhood . In J. Cassidy, ed. and P.R. Shaver (eds), *Handbook of*

Attachment: Theory, Research, and Clinical Applications , 2nd Edition, pp. 192–216.
New York: Guilford Press.

Weiss, R.S. (1973) Loneliness: The Experience of Emotional and Social Isolation .
Cambridge, MA: MIT Press.

<http://dx.doi.org/10.4135/9781446249222.n34>

Handbook of Theories of Social Psychology

Shared-Reality Theory

Contributors: Gerald Echterhoff

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Shared-Reality Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n35>

Print pages: 180-200

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n35>

[p. 180 ↓]

Chapter 35: Shared-Reality Theory

GeraldEchterhoff

Abstract

In providing an overview of shared reality theory, the chapter first presents the current formulation of the theory, with a focus on assumptions about critical conditions for the occurrence of shared reality and its underlying psychological processes. The second section provides a selective review of empirical evidence, focusing on interpersonal communication as a main arena for shared reality creation and motivational underpinnings. In the third section, the intellectual history of shared reality theory is outlined. Based on the original contributions by Tory Higgins and his colleagues, the ancestry of the theory is traced to the domain of language and communication, phenomenological approaches in sociology and philosophy, and social influence research. The final section provides illustrations of the applicability of the theory, and a discussion of the theory's utility.

Introduction

Creating shared views about the world is ubiquitous. For instance, when people meet a new employee at their workplace, they tend to form their impressions of the newcomer jointly with their colleagues, and they feel more confident in their impressions when others agree. People take into account the views of others, especially significant others (see Andersen and Chen, 2002), to appraise experiences and events, and to construct or verify views about various types of issues (Hardin and Higgins, 1996). Social sharing allows us, for example, to evaluate other people or groups; to form general political, moral, or religious convictions; and even to develop and maintain a sense of who we are and what we want (Higgins, 1996b; James, 1890; Sullivan, 1953). The absence of

social sharing can have detrimental consequences for people's wellbeing, their feelings of connectedness, and sense of reality. When interaction partners withhold an expected shared reality, such as in the classical conformity studies by Asch (1951), people are left uncertain, uncomfortable, even physically agitated.

Accounts emphasizing the social underpinnings and interpersonal nature of our representations of reality have circulated, in varying shapes and forms, for a long time in the social sciences and psychology (e.g., Asch, 1952; Bar-Tal, 1990, 2000; Cooley, 1964; Festinger, 1950; Heider, 1958; Lewin, 1947; Mead, 1934; Merton and Kitt, 1950; Moscovici, 1981; Newcomb, 1959; Resnick et al., 1991; Rommetveit, 1974; Schachter, 1959; Schütz, 1967; Sherif, 1935, 1936). [p. 181 ↓] Given the long history of these approaches and compared with most theories covered in this handbook, shared reality theory as a distinct theory of social influence on reality construction is relatively young. While the conceptual framework has been developed mainly over the last 20 years (Hardin and Higgins, 1996; Higgins, 1992, 1999), empirical work informed by the theory has been mostly published only within the past five years (for a review, see Echterhoff et al., 2009a).

The theory's concern with the construal of reality conjures up a classical issue that has concerned scholars in philosophy, cognitive psychology, and the neurosciences for many decades. How can people, given the highly constructive and self-contained operation of the human mind, distinguish between what is real and what is mere imagination, dream, or fantasy? Apparently, humans are equipped with cognitive and neural mechanisms, such as comparing predicted and actual sensory input, that tell them whether the output of mental operations and their mental models are sufficiently consistent with external stimulus conditions, that is, whether they are in touch with the real world (e.g., Frith, 2007; Johnson and Raye, 1981). In contrast, shared reality theory focuses on processes that are different from or subsequent to such low-level monitoring of reality. For instance, when members of a work team meet a new colleague, they try to find out what kind of person the newcomer is, whether she is trustworthy, sociable, and open-minded. This allows them to evaluate the newcomer, to predict her actions, and to interact with her purposefully. The relevant question here is not whether the observed events are real, that is, whether the team members trust their perception of the newcomer's appearance and behavior as real (versus imagined). Rather, the question refers to the attributes and qualities of the target entity and the meaning of the

observed events; that is, how the team members think about, categorize, judge, and evaluate the newcomer based on their perceptions and observations (see Higgins, in press).

Whereas the cognitive factors of reality construction, like the accessibility of knowledge and its semantic applicability, have been studied extensively and intensively by social psychologists in past decades (see, e.g., Higgins, 1996a, in press), genuinely social factors have typically received less attention. In the example, the experienced reality or truth about the newcomer can result from cognitive processes like the activation and application of pertinent knowledge, but it can also result from sharing impressions about the newcomer between the old team members (Levine and Higgins, 2001). For instance, the team members may create through conversation a shared view as to whether the talkativeness of the newcomer means that she is sociable or cordial (positive traits), rather than conceited or ingratiating (negative traits).

Nowadays, most psychologists would probably accept, or at least not actively dispute, the general notion of a social foundation of mental representations of the world. However, it is less common for psychologists, given the long-standing individual-centered orientation of the discipline, to make this issue the subject of deeper analysis and empirical investigation in its own right. The latter projects represent precisely the agenda of shared reality research. Given the occasional generality and diversity of earlier related theorizing, a key challenge for theory construction was to steer the concept of shared reality toward a greater, adequate level of specificity and carve out its distinctive and unique potential. A challenge for empirical research was to capture the novel theoretical assumptions and hypotheses in experiments that would convince the peer community, and to assess the occurrence – that is, people's experience – of shared reality empirically. The present chapter provides an overview of the results and some of the history of this endeavor.

The first section of this chapter is devoted to the conceptualization of shared reality and assumptions about the critical conditions for the occurrence of shared reality, along with [p. 182 ↓] the distinct psychological mechanisms that underlie it. In the second section, I will briefly review empirical evidence supporting the theory, which focuses on shared reality created in interpersonal communication and its motivational underpinnings. In the third section, the intellectual history of shared reality theory is outlined, including the

original contributions by Tory Higgins and his colleagues; the ancestry of the theory is traced to the domains of language and communication, phenomenological approaches in sociology and philosophy, and social influence research. The final section provides illustrations of the applicability and utility of the theory.

Shared Reality Theory: Conceptualization and Critical Conditions

According to the most current proposal (Echterhoff et al., 2009a), shared reality is defined as the product of the motivated process of experiencing with others a commonality of inner states about the world. This conceptualization presumes that four main conditions underlie shared reality. First, the commonality between individuals that is implied by a shared reality refers to their inner states and not just their overt behaviors. Second, shared reality is “about something” – that is, it implies a target referent about which people create a shared reality. Third, shared reality as a product cannot be divorced from the process through which it is attained – in particular, the underlying motives. Fourth, there is no shared reality unless people experience a successful connection to someone else's inner state. These conditions will now be elaborated in turn.

According to the first condition, a shared reality involves a commonality between people's inner states, which include their beliefs, judgments, feelings, or evaluations concerning a target referent. To achieve a shared reality, people cannot simply replicate the observable behavior of others; instead, they need to obtain a sense of others' inner states about the world. For the occurrence of a shared reality, a correspondence between externally observable states or behaviors is not sufficient – it needs to involve a commonality between inner states (see Brickman, 1978). This claim is supported by the fundamental and well-established role that the perception of others' inner states plays in human development, motivation, and sociality. People know not only that the outcomes for a person (self or other) depend critically on another person's overt responses to that person (e.g., Ostrom, 1984), but also that the other person's responses are mediated by his or her inner states, such as his or her attitudes and beliefs (Higgins, 2005, 2010). Indeed, the discovery of the mediating role of others'

inner states in how they respond to the world is a significant step in human development (see Higgins and Pittman, 2008). Once this level of social consciousness is reached, others' inner states begin to play a vital role in human self-regulation (Higgins, 2010).

The achievement of the first condition requires processes that allow people to pick up or infer someone else's inner state. Psychological research suggests a plethora of mechanisms by which this can be accomplished (see, e.g., Higgins and Pittman, 2008; Malle and Hodges, 2005). For instance, people draw on various aspects of others' nonverbal behavior, such as their facial expressions and gestures, to intuit the others' feelings, needs, and intentions. They grasp others' mental states, such as others' beliefs, attitudes, and feelings, drawing on mechanisms like conscious reasoning, unconscious simulation, and theory of mind (e.g., Keysers and Gazzola, 2007); causal theories and schemata (e.g., Heider, 1958; Malle, 1999); or projection of their own inner states (e.g., Keysar and Barr, 2002; Nickerson, 2001). This precondition is a building block of the first condition for shared reality – the perceived sharing of inner states and not just overt behaviors.

The previous argument makes implicit reference to the second condition of shared [p. 183 ↓] reality – shared reality is about some target referent. For a shared reality to occur it is not sufficient that people simply exhibit corresponding inner states, such as corresponding heart rates or mood states. If corresponding inner states are not about (i.e., not in reference to) some aspect of the world, then one cannot speak of a shared reality. This is because *reality* refers to the objects or referents of knowledge – that is, to phenomena that are experienced by actors as being part of the world in the present, as well as in the past and future (such as future desired end-states; Higgins and Pittman, 2008). Thus, shared reality goes beyond simply replicating another person's inner state in that it requires sharing states that are *about* some target referent: for example, about a new colleague at work, about a specific TV program, about a particular politician, or about abstract political or religious issues (see Jost et al., 2007).

Like the first condition, this second condition – that shared reality is *about* some target referent – requires that a critical precondition be met. Specifically, it requires mechanisms that allow people to infer the target referent of their sharing partner's inner state, such as the referent of another person's feeling. Research has identified various mechanisms by which this can be achieved. One basic mechanism is to follow the

direction of someone else's eye gaze (e.g., Baron-Cohen, 1995; Tomasello et al., 2005; Tomasello, 2008) to identify the referent of that person's sustained interest or emotional response, such as what it is that she or he fears. Eye-gaze following, together with imputing intentionality to the other person, allows the allocation of shared interest in an object (Baron-Cohen, 1995). Other mechanisms include following someone else's pointing movements or manipulations of objects (Clark, 2003; Tomasello, 2008) and interpreting verbal utterances as referring to an object (Clark, 1996; Clark and Marshall, 1981).

According to the third condition, the occurrence of shared reality depends on the motives that drive the achievement of common inner states. By this view, what needs to be taken into account is the source, or process history, of a commonality of people's inner states. The sheer fact or presence of a commonality is not sufficient. An analogy would be that democracy concerns not only consensus as an outcome or state of agreement but also the processes by which people reach a consensus (see, e.g., Bohman and Rehg, 1997). How a consensus or agreement is reached and whether the right procedures are observed to arrive at a consensus are, in many cases, more important than the product or outcome itself (Mackie and Skelly, 1994). More generally, end states often attain their value from how they were reached and not just from the outcome per se (Higgins, 2006).

What, then, are the motives driving the creation of a shared reality? Two motives have figured prominently in the literature on social motivation in general and on shared reality in particular: epistemic and relational motives (Bar-Tal, 2000; Fiske, 2007; Hardin and Conley, 2001; Hardin and Higgins, 1996; Jost et al., 2007). *Epistemic motives* refer to the need to achieve a valid and reliable understanding of the world (Hardin and Higgins, 1996) and to establish what is real (Higgins, *in press*). Humans are motivated by what Bartlett (1932) called effort after meaning, a fundamental need to understand the events and circumstances of their lives (e.g., Kagan, 1972), to extend their knowledge continuously (Loewenstein, 1994), and, no less important, to experience themselves as successful in this endeavor (Higgins, *in press*). Achieving such epistemic goals has various beneficial consequences; for example, a sufficiently accurate understanding of the world allows humans to operate successfully in their environment. It ranges from basic knowledge about how to survive by securing necessary resources and avoid vital risks to sophisticated mental models of how to create and maintain important

relationships or to develop one's professional career. The strength of epistemic motives typically increases with uncertainty or ambiguity about a target referent (e.g., Berlyne, 1962; Hogg, 2007). [p. 184 ↓] Consistent with this notion, Festinger (1950) argued that the more ambiguous and difficult to interpret experiences are, the more people seek a social reality provided by appropriate (i.e., sufficiently trustworthy) others (see also Byrne and Clore, 1967; Deutsch and Gerard, 1955; Sherif, 1936). Given that shared reality is about a target referent, it follows that the creation of shared reality always serves, at least to some extent, epistemic motives.

Relational motives induce people to affiliate and feel connected to others. Feeling connected to others has several positive consequences, including emotional wellbeing, a sense of security, and self-esteem (e.g., Baumeister and Leary, 1995; Diener and Seligman, 2002). The desire for connectedness is reflected, for example, in the affiliative tendency that people exhibit when they are confronted with potentially anxiety-arousing situations (Schachter, 1959).

The motivational process that has led to a commonality of inner states is assumed to be critical to whether the commonality is a shared reality. Adopting another person's inner state, for example, could be driven by instrumental goals of securing beneficial social responses or maximizing personal outcomes (see Higgins, 1981; Jones and Thibaut, 1958). Social actors pursue such instrumental goals, for instance, when they ingratiate themselves with others (Jones, 1964) or take the perspective of a competitor to prevail in a social conflict (Epley et al., 2006). In such cases, actors adopt another person's inner state not because they want to achieve a better understanding of a reference target or to establish what is real, but because they hope to attain other, ulterior goals. In such cases, the adopted commonality is not a shared reality.

According to the fourth condition, a shared reality requires that the participating individuals actually experience the sharing – that is, that they experience a commonality with someone else's inner state. Consistent with this view, Bar-Tal (2000) has argued that sharing of beliefs entails more than merely an objective commonality between people that can be identified by an external observer. Instead, sharing must involve the subjective experience or awareness of a commonality. Even if people are motivated to share inner states with others, they may end up not establishing an experienced commonality. Thus, it is not enough to have taken action to create a commonality.

with another person's inner state in the service of appropriate (such as relational or epistemic) motives. It is also necessary that one perceive the commonality to have been, in fact, established.

By including the fourth condition that individuals actually experience the sharing, Echterhoff et al. (2009a) emphasized the critical role of the subjective sense of sharing. This aspect can be further elaborated in the context of the first and second condition for shared reality, thus suggesting possible interrelations among the conditions. For the achievement of shared reality, people need to subjectively experience both the commonality of inner states and the referential aboutness of inner states. From this perspective, there can be a shared reality even if both assumptions of sharing are objectively wrong. That is, for Person A to experience a shared reality with Person B, it is not necessary for B to actually have the same inner state as A or for B's inner state to actually refer to the same referent that A has in mind. What is critical is that A believes that B's inner state and the referent of that inner state match A's inner state and referent. Consider, for example, a new member (A) in a research lab who believes that the current members in the lab are arrogant and wants to create a shared reality with another newcomer (B) about these members. For Newcomer A to have a shared reality with B, it is critical that A infer (e.g., by observing that B acts in a tense and uncomfortable manner at a lab meeting) that B has a shared inner state about the current members as referent (i.e., B also believes that they are arrogant). If Newcomer B later makes clear that the current members are her academic idols and that she always feels uncomfortable in encounters with admired people, then A's sense of a shared reality will be eliminated.

[p. 185 ↓]

The unique contribution of this concept of shared reality can be seen exactly in the formulation of the four critical conditions and its building blocks. This addition to the theory has enhanced the precision, integration, and testability of the theory, and it affords clear distinctions between shared reality and related concepts such as common ground, empathy, perspective-taking, embodied synchrony (see Semin, 2007), and socially distributed knowledge (for a discussion, see Echterhoff et al., 2009a).

Empirical Evidence Supporting the Theory

In the following, I review empirical studies that have examined the creation of shared reality in interpersonal communication, which is arguably the main vehicle for social sharing (Berger and Luckmann, 1966; Higgins, 1992; Higgins and Rholes, 1978). When people are motivated to create a shared reality with others, they often communicate with these others about a target referent. The communication studies presented here are based on the hypothesis that communicating about a target referent can affect communicators' cognitive representations of that target. Studies employing the "saying-is-believing" paradigm have demonstrated such communication effects on subsequent cognition (e.g., Higgins and Rholes, 1978; Higgins et al., 2007; Sedikides, 1990; for reviews see Higgins, 1992, 1999; McCann and Higgins, 1992). In this paradigm, participants are introduced to an ostensible referential communication task (involving a communicator, a target, and an audience) in which they take the role of the communicator. The participants, who are typically students, read an essay about another student (the target person) who supposedly has volunteered to be part of a long-term research project on interpersonal perception. They are told that their task is to describe the target person's behaviors – without mentioning the target's name – to another volunteer (the audience) who knows the target person. On the basis of their message description, the "audience" volunteer would try to identify the target person as the referent of the message from among a set of several possible targets in the alleged research project.

A short essay consisting of several passages provides the input information about the target person. The behaviors described in each passage are evaluatively ambiguous; they can be interpreted as indicating either a positive or a negative trait with approximately equal likelihood (e.g., "persistent" versus "stubborn" or "independent" versus "aloof"). For example, the behavior described in the following sample passage could be labeled as either "independent" or "aloof": "Other than business engagements, Michael's contacts with people are surprisingly limited. He feels he doesn't really need to rely on anyone." (e.g., Echterhoff et al., 2008). To manipulate the audience's supposed attitude toward the target person, the researchers informed the participants (in an offhand way) that their audience either likes the target (positive audience attitude)

or dislikes the target (negative audience attitude). In their subsequent communication, participants typically exhibit audience tuning: They evaluatively tailor, or “tune,” their messages to their audience’s attitude (i.e., they create evaluatively positive messages for an audience who likes the target and evaluatively negative messages for an audience who dislikes the target).

After a delay (from approximately ten minutes in some studies to several weeks in other studies), researchers test the participants’ memory for the original input information. Participants are asked to recall, as accurately as possible, the original essay about the target person in a free, written format. In demonstrations of the saying-is-believing effect, the evaluative tone of the communicators’ own recall for the original input information matches the evaluative tone of their previous, audience-tuned message. In other words, communicators’ own memory representations of the message topic [p. 186 ↓] reflect the audience-tuned view expressed in their message rather than just the original target information. Communicators end up believing and remembering what they said rather than what they originally learned about the target.

After the initial demonstrations of this “saying-is-believing” effect (e.g., Higgins and Rholes, 1978; Higgins and McCann, 1984), a number of studies using other paradigms have shown that people’s mental representations of an experience can be profoundly shaped by how they verbally describe the experience to others (e.g., Adaval and Wyer, 2004; Tversky and Marsh, 2000; for reviews, see Chiu et al., 1998; Marsh, 2007). Thus, the influence of verbal communication on subsequent cognition is well established. Also, the saying-is-believing effect in particular has been replicated with several variations in methodology and extended to new areas. For instance, although the effect was originally demonstrated for tuning to the audience’s attitude toward the target (Higgins and Rholes, 1978), it has also been found for tuning to the audience’s knowledge about the target (Higgins et al., 1982). Also, the effect occurs regardless of whether communicators know their audience’s view before or after encoding the input information (Kopietz et al., 2010, Experiment 1). The effect has been extended from situations in which the communication topic is a single individual to situations in which the topic target is a small group (Hausmann et al., 2008). Furthermore, the effect occurs not only with verbal stimulus material as input information about a target, but also with complex visual input material, namely video-filmed behaviors of target persons (Hellmann et al., in press; Kopietz et al., 2009).

Several studies demonstrated that the saying-is-believing effect occurs to the extent that communicators create a shared reality with their audience about the target person, as characterized by the four conditions outlined earlier (e.g., Echterhoff et al., 2005, 2008; Echterhoff et al., 2009b; Hellmann et al., in press; Higgins et al., 2007; Kopietz et al., 2009; Kopietz et al., 2010). In the studies by Echterhoff, Higgins, and colleagues, for example, communicators' memory representations of the target person (assessed by free recall) were biased by their audience tuning under conditions that support creating a shared reality but not under conditions that undermine creating a shared reality. The creation of a shared reality can fail when any one of the four conditions described earlier fails to be sufficiently satisfied.

The bulk of extant empirical evidence relates to the third condition of shared reality; that is, the motivation underlying creating an interpersonal commonality, and thus this review emphasizes findings relevant to this condition. A discussion of findings regarding the other three conditions can be found in Echterhoff et al. (2009a), who also review evidence that rules out alternative explanations for the saying-is-believing effect such as differential reduction of cognitive dissonance (Festinger, 1957), self-perception (Bem, 1967), and source confusion. What the research shows is that communicators create a shared reality with their audience only when their production of audience-congruent messages is appropriately motivated; that is, the motivation behind the creation of commonality with another person's inner state is critical for a shared reality.

In one set of studies, Echterhoff et al. (2008) directly manipulated the goals underlying audience tuning. It was assumed that, in the standard saying-is-believing conditions (e.g., Higgins and Rholes, 1978), audience tuning serves epistemic motives that are characteristic of shared reality. Specifically, the evaluative ambiguity inherent in the behavioral-input information about the target person should elicit the epistemic motivation to reduce uncertainty. By tuning messages to the audience, communicators construct an audience-congruent representation of the target and thus attain a greater sense of certainty about what the target is like. In the Echterhoff et al. (2008) studies, this standard "shared reality-goal" condition was compared with conditions in which audience tuning served nonshared [p. 187 ↓] reality goals. The nonshared reality goals included obtaining monetary incentives for producing an audience-congruent message and entertaining the audience with an exaggerated, caricature-like description of the target person (Echterhoff et al., 2008, Experiments 2a and 2b). Based on the

above rationale, it was hypothesized that communicators in the shared reality-goal condition should adopt the audience's inner state during message production to reduce uncertainty about the target person. In contrast, in the nonshared reality-goal conditions communicators should adopt their audience's inner state primarily to attain goals unrelated to the epistemic motivation that is characteristic of shared reality; they pursue alternative, or "ulterior," goals that are not conducive to a shared reality.

As predicted, it was found that communicators in these alternative, nonshared reality-goal conditions tuned their messages even more strongly to their audience's attitude than did communicators in the shared reality-goal condition (for the sake of the incentive or entertainment), but, nonetheless, the audience-tuning memory bias was not found. In contrast, the memory bias was found as usual in the standard shared reality-goal condition. Consistent with shared reality assumptions, additional measures revealed that audience tuning was experienced as being motivated by external demands to a greater extent in the alternative-goal conditions than it was in the shared reality-goal condition. Also, communicators' *epistemic trust* in the audience and in their audience-congruent message was significantly higher in the shared reality-goal condition than in the alternative nonshared reality-goal conditions.

These findings suggest that when people merely want to go along with another person – for instance, to obtain rewards from this person – a shared reality with that person is not produced. When people generate representations corresponding to another person's inner state without being motivated to create a shared view about a target, they do not achieve a shared reality. What matters is not the fact of a commonality with another person per se, but the motivation that produces the commonality.

In another set of studies, the creation of a shared reality was shown to depend on whether communicators were or were not motivated to share inner states with the particular person who was the audience for their message (Echterhoff et al., 2005, Experiment 2; Echterhoff et al., 2008, Experiment 1; Kopietz et al., 2010). Presumably, communicators do not regard just any person to be an appropriate partner with whom to share inner states. As suggested by research on social comparison and group-anchored knowledge (e.g., Festinger, 1950; Kruglanski et al., 2006), individuals regard others who possess certain qualities, such as sufficient similarity and trustworthiness, as more appropriate partners with whom to share reality than others who lack these

qualities. Among these qualities, membership in a perceiver's ingroup (versus outgroup) is particularly important.

As suggested by various strands of intergroup research (e.g., Hogg, 2007; Kruglanski et al., 2006; Levine and Higgins, 2001), people should be less motivated to create a shared reality with outgroup members than with ingroup members. Nonetheless, in the standard saying-is-believing paradigm, which involves a referential communication task, communicator participants can still be expected to tune their message to an outgroup audience. However, compared to tuning messages to an ingroup audience, tuning messages to an outgroup audience should be motivated more by task fulfillment and politeness demands than by the desire to achieve a shared reality with the audience for epistemic and relational motives. In the standard saying-is-believing paradigm, shared reality motives are typically induced, but when the audience is an outgroup member alternative, nonshared reality motives should take precedence. Thus, if the motivation behind audience tuning is critical, then communicators tuning messages to an outgroup audience should exhibit little if any audience-congruent recall bias.

[p. 188 ↓]

These predictions were borne out in studies by Echterhoff, Higgins, and colleagues. Although messages were tuned to both outgroup audience and ingroup audience, communicators talking to an outgroup audience did not incorporate the audience-tuned message into their own memory of the target. They also exhibited lower epistemic trust in their audience's view than did communicators tuning to an ingroup audience (Echterhoff et al., 2005, Experiment 2; Echterhoff et al., 2008, Experiment 1). Furthermore, participants (German students at a German university) who communicated to an audience belonging to a stigmatized outgroup (Turks) reported more often that they made an active effort to adapt their messages to their audience's views than did participants communicating to an ingroup (German) audience (Echterhoff et al., 2008, Experiment 1). These findings suggest that people producing audience-congruent messages merely to comply with external demands (e.g., behaving in a polite or unprejudiced manner; see Dovidio et al., 2002; Richeson and Trawalter, 2005) do not create a shared reality.

While the previous review has focused on the effects of shared reality driven communication on the communicators themselves, these effects are likely to apply to the recipients as well. Support for this notion comes from a recent study by Stukas et al. (2010). These authors found that recipients' beliefs about a target group were biased by messages tuned toward their presumable beliefs, whereas in fact they initially did not know the target group and thus did not hold preexisting attitudes. Consistent with shared reality theory, the effect on recipients was stronger for those recipients who reported a greater willingness and experience of shared reality with the communicator. Thus, communicators' expectations about recipients' attitudes toward a group can initiate a confirmatory process by which both communicators and recipients come to hold the expected attitudes, thus giving rise to a shared reality created without anyone holding an initial view of that "reality"!

Intellectual History of the Theory

The social foundation of basic psychological phenomena has been conceptualized and discussed for a long time in different disciplines, including social psychology (e.g., Asch, 1952; Bar-Tal, 1990; Festinger, 1950; Hardin and Higgins, 1996; Heider, 1958; Levine and Higgins, 2001; Lewin, 1947; Moscovici, 1981; Schachter, 1959; Sherif, 1935, 1936), the general social sciences and sociology (Cooley, 1964; Mead, 1934; Schütz, 1967; Thompson and Fine, 1999), memory and cognition (Graf et al., 2010; Hirst and Echterhoff, 2008; Hirst and Manier, 2002; Marsh, 2010; Smith and Semin, 2004; Weldon, 2001), psycholinguistics (Pickering and Garrod, 2004), organizational behavior (Salas and Fiore, 2004), developmental psychology (Meltzoff and Decety, 2003), evolutionary psychology (Caporael, 2007, 2010; de Waal, 2008), social neuroscience (e.g., Gallese et al., 2004; Iacoboni, 2008, in press), biology (e.g., Dunbar and Shultz, 2007), and philosophy (e.g., Thagard, 1997). While this field spans a wide range of approaches with greatly varying terminology, the development of the distinctive concept of shared reality can be attributed and dated much more specifically. It was originally achieved and literally "nurtured" from its very inception in the 1990s to the present day by Tory Higgins. He, joined over the years by collaborators, both laid the groundwork and continuously developed the theory over the years. Hence, the following brief history

is oriented by the publications of Tory Higgins and his colleagues during each phase of the development of shared reality theory.

These contributions – primarily Higgins (1992), Hardin and Higgins (1996), Higgins (1999), Echterhoff et al. (2009a), but to some extent already Higgins (1981) – contain not only the substance of the concept of shared reality but also review the scholarly works that have served as input to the theory. Taken together, the various precursors and sources [p. 189 ↓] of inspiration can be organized into three main areas: interpersonal communication and language use, phenomenological approaches in philosophy and sociology, and social influence research. Intermittently, the personal dimension of the theory development is illustrated, including biographical circumstances.

Let me begin with the first and probably most important thread in the theory's lineage. Throughout the development of the theory, a key influence and primary domain for empirical investigation has been interpersonal communication. Higgins' concept of communication as an *interpersonal game* has been the backdrop of the explicit introduction of shared reality theory in a paper published, quite fittingly, in the *Journal of Language and Social Psychology* (Higgins, 1992). It was also the topic of an early chapter, "The 'Communication Game,'" that foreshadowed the subsequent, explicit formulations (Higgins, 1981). His theorizing was inspired by pragmatic approaches to language (e.g., Austin, 1962; Grice, 1975; Rommetveit, 1974) and, particularly, Wittgenstein's (1953) characterization of verbal communication as a game. From the perspective of pragmatics, language is a motivated, context-dependent means of interpersonal communication, following (explicit and implicit) rules and assumptions. Motives of communication include conveying information or meaning, creating or maintaining a social relationship, achieving a shared understanding with others, but also influencing others' behavior, maximizing beneficial social responses, and accomplishing a joint task.

In Higgins (1981) and Higgins (1992), communication is characterized, *inter alia*, as a social action that involves shared and context-dependent expectations and rules concerning the interlocutors' roles and appropriate language, and that requires the mutual consideration of each other's characteristics, specifically their knowledge, attitudes, and intentions. From these and other assumptions about language use in

social interaction, Tory Higgins derives several rules of the communication game, including some that are at the heart of empirical work on audience-tuning effects in the saying-is-believing paradigm. For instance, communicators should say what is relevant, and give neither too much nor too little information (Grice, 1975), and take into account the audience's perspective, knowledge, attitudes, and preferences in their language use. The latter rule gives rise to audience tuning in communicators' message production. As described in the previous section, audience tuning in interpersonal communication has represented the principal arena for empirical demonstrations of shared reality processes in the saying-is-believing paradigm.

One section of the early chapter refers to the "sharing of a social bond" and "social reality" (Higgins, 1981) and contains elements that were elaborated later in the explicit formulations of shared reality theory. In particular, the early chapter foreshadows the key role of goals and motives for the creation of shared reality. The "convergence of opinions and judgments" (Higgins, 1981: 376) – which in the current terminology represents a commonality of inner states – is assumed to depend on whether the interlocutors have "a desire to maintain the social bond and share a common definition of social reality" (Higgins, 1981: 376). These motives translate quite seamlessly into the two main types of motivation, affiliative and epistemic, that are assumed to drive shared reality according to the recent theory formulation (see second section).

In both the 1981 and the 1992 paper, Higgins illustrated the potential and validity of the communication-game approach to a large extent with empirical findings from saying-is-believing and closely related studies. However, the first saying-is-believing publication (Higgins and Rholes, 1978) and empirical follow-up papers were framed quite differently, mostly in information-processing terms based on the nascent social cognition approach, which began to fascinate increasing numbers of social psychologists in the 1970s. It would take almost a decade for Higgins to return to his communication game analysis to account for the saying-is-believing effect.

[p. 190 ↓]

Ironically, at the time he began the saying-is-believing research, a key interest of Tory Higgins, who had been a joint honors anthropology–sociology student at McGill University in Montreal, was the inter-relation of language, thought, and society, and

especially the linguistic relativity hypothesis. The thrust and inspiration of the 1981 chapter, particularly its focus on the pragmatics of language and communication, is chiefly owed to that interest. The irony is that this interest originally emphasized the interpersonal and society-related aspects of the interrelation, as reflected in the “communication game” label. However, when the saying-is-believing research formally appeared as a published article, it was instead the language-and-thought aspect of the inter-relation that was highlighted. This framing matched the social cognition emphasis during that point in history, not to mention the social cognition emphasis of the editor of the journal in which the first article appeared (Bob Wyer).

The core idea of the linguistic relativity hypothesis is that language provides the essence of human thinking, and that mental representations entertained in our minds are inextricably linguistic. The position has been epitomized by the Sapir–Whorf hypothesis, named after the anthropologist Edward Sapir and his student Benjamin Lee Whorf. In its strong version, this hypothesis states that our experiences with the world are intrinsically linguistic and that cognition is determined by the thinker's language (see Hunt and Agnoli, 1991). Both Sapir (1964) and his student Whorf argued that differences in the structure of speakers' language create differences in cognition: “We dissect nature along the lines laid down by our native languages” (Whorf, 1956: 213). Since the first budding of interest in the 1950s, the hypothesis provoked controversial debates and stimulated a substantial corpus of research, particularly on cognitive differences between speakers belonging to different language communities, such as speakers of Mandarin, English, or Navajo.

Research in this field addresses primarily the extent to which the structural aspects of a language affect the speakers' cognition. By this view, language is predominantly treated as an underlying structure, much in the sense of what Chomsky called linguistic competence. Also, the language-and-thought debate traditionally focused on effects of language at the lexical level, such as whether memory for color stimuli depends on the availability (versus lack) of certain color terms in a language. However, Tory Higgins was interested – as was his dissertation supervisor at Columbia University, Bob Krauss (e.g., Chiu et al., 1998) – identifying effects of the pragmatic usage of language on cognition. According to this idea, which was a major inspiration for the saying-is-believing work, language may influence thought not so much at the structural or lexical level, but at the level of actual language use in motivated, rule-based interpersonal

communication. This influence would be due not to the type of language, but concrete instances, or tokens, of verbal communication (Holtgraves and Kashima, 2008). As described above, tuning a message to one's audience's characteristics (attitude, knowledge) is a central phenomenon exemplifying the pragmatic use of language in verbal communication.

The second line of theoretical precursors can be traced to phenomenological approaches in philosophy and sociology. Related to the second condition of shared reality outlined above, philosophers like Husserl (1931) and Brentano (1974), suggested that directedness, or "aboutness," is a general characteristic of human thinking. This understanding is consistent with the social-psychological notion, also emphasized by shared reality theory, that people want to increase their knowledge of the world and hence represent their own and others' behavioral responses as being *about* something (see Heider, 1958). This notion draws attention to the triadic relation implied by many formulations of shared reality theory, specifically the relation between one person experiencing sharing, another person (a "sharing partner") or group [p. 191 ↓] of persons with whom the sharing is experienced, and a target referent of the sharing (cf. Tomasello et al., 2005).

As outlined in the second section, shared reality permits a perceiver to experience some target referent in common with another person. There are other cases of social sharing that do not meet this condition. Phenomena such as empathy (de Waal, 2008) and emotional and mood contagion (Neumann and Strack, 2000) do not require that the perceiver share the other person's view about a target referent. The importance of directedness, or aboutness, for shared reality was emphasized first in Higgins (1999). The basis for the inclusion of *aboutness* was a paper published shortly before (Higgins, 1998), in which the importance of aboutness as a general principle of human inference and judgment was emphasized.

The notion that people's *experience* of reality is socially established resonates with earlier conceptualizations in phenomenological sociology, particularly the sociology of knowledge (Berger and Luckmann, 1966; Garfinkel, 1967; Schütz, 1967). For instance, Berger and Luckmann argued that "[t]he reality of everyday life ... presents itself to me as an intersubjective world, a world that I share with others." (1966: 23) Scholars in these fields also understood that social actors "are motivated to create a sense, even an

illusory sense, that they share a common universe,” so that they might “generate a tacit presumption that there is an external factual order ‘out there’” (J.H. Turner, 1987: 19). This emphasis of the motivational underpinnings of a shared world view is compatible with the third condition formulated in shared reality theory, which holds that regarding the commonality of inner states only as an outcome or end product would overlook important psychological underpinnings, specifically the underlying motives and the experience of sharing. Furthermore, the earlier sociologists realized that while people do not have direct access to each others’ inner states, they can still “put themselves in each others’ place” (Turner, 1987: 18) by means of interpersonal practices such as exchanging and interpreting signs. Such practices, it was assumed, produce the subjective experience of successfully connecting to others’ inner states.

The inspirations in this second ancestry line gained prominence, and were cited more comprehensively, as shared reality theory evolved over time and attracted coauthors. The work of scholars at the interface of sociology and anthropology (e.g., Harold Garfinkel) were cited already in the earlier publications (Higgins, 1981, 1992). However, with the contributions of subsequent close collaborators, to precursor approaches in the humanities, particularly philosophy and sociology, were increasingly acknowledged, for instance in the first comprehensive account of shared reality (Hardin and Higgins, 1996) and other subsequent developments of the theory (Echterhoff et al., 2009a).

Research on social influence represents the third line of intellectual inspiration. Social influence occurs when an individual’s responses like behaviors, attitudes, and judgments are shifted to become consistent with the position of one or more others as a result of contact or interaction with these others. A prototypical instance of social influence is persuasion. One possible result of social influence is a commonality of *inner states*, such as attitudes and judgments about something, between at least two people. The close relation between social influence and shared reality is conspicuous in most publications on shared reality (for a direct empirical investigation, see Pinel et al., 2010). The early chapter by Higgins (1981) discussed implications of the communication-game concept for social influence, specifically the role of language use in persuasion. Also, in the first comprehensive exposition of the theory Curtis Hardin and Tory Higgins (1996) discussed how social influence in groups (e.g., Asch, 1951; Deutsch and Gerard, 1955; Festinger, 1950; Latané and Wolf, 1981) can be interpreted through the lens of shared reality theory (also see Levine and Higgins, 2001). In Echterhoff et al. (2009a), a section

is devoted to the relation between the concept [p. 192 ↓] of shared reality and research on social influence in groups and work teams.

My own involvement in research on shared reality in interpersonal communication was mediated and facilitated to a large extent by my interest in effects of the social and communicative context on memory and remembering (e.g., Echterhoff et al., 2007; Echterhoff and Hirst, 2009; Lindner et al., 2010). It began in the years 2001/2002 when I worked as a postdoctoral researcher with Bob Krauss and Tory at Columbia University. My doctoral dissertation had been supervised by Bill Hirst, an eminent memory researcher at the New School for Social Research who became increasingly interested in the role of the social context for memory processes (e.g., Cuc et al., 2007; Hirst and Manier, 2002). This orientation contributed to the focus on memory, specifically the audience-congruent recall bias, as the main dependent variable in saying-is-believing studies of shared reality, which I embarked on together with Tory Higgins during my postdoc research. Forming one of the rare intersections of personal and global history, I almost gave up completing the first experiment, conducted with the assistance of my doctoral student Stephan Groll, in the immediate aftermath of the shocking events of September 11, 2001, in New York City – we were petrified ourselves, and study participants ceased showing up in the lab.

The studies were initially motivated by the search for a genuinely social, perhaps shared reality, account of the saying-is-believing effect that would go beyond the conclusions of the pioneering studies from around the 1980s (see McCann and Higgins, 1992). Given my own background at the time, the saying-is-believing effect struck me as a type of social influence that differs from “classical” social influence as it affects the source rather than the recipient of communication and represents a subtle, self-produced bias (see Echterhoff and Hirst, 2009). For a long time, memory researchers had paid little attention to genuinely social influences on memory. Conversely, social psychologists had focused on attitude, judgment, and behavior rather than memory as the object of social influence (see Bless et al., 2001).

Applicability and Utility of Shared Reality Theory

Shared reality theory addresses fundamental psychological questions such as how people establish what is real and satisfy basic epistemic needs (see Higgins, in press). Thus, the theory is potentially applicable to various phenomena involving the establishment and experience of reality (for constraints, see the introduction). In a general sense, the goal of creating a shared reality plays a critical role in how communication contributes to the formation and maintenance of people's knowledge. More specifically, the creation of shared reality in social interaction is relevant and applicable to a host of real-life issues, such as persuasion of self and others through political and religious speeches (Vedantam, 2008), belief justification in close relationships and families (Jost et al., 2007; Magee and Hardin, 2010), stereotyping and prejudice (see Huntsinger and Sinclair, 2010; Kashima et al., 2010; Klein et al., 2010; Sinclair et al., 2005a, 2005b), interethnic interaction (Conley et al., 2010), decision-making and performance in work teams (see Salas and Fiore, 2004), the transmission of culture (Echterhoff and Higgins, 2010), and the protection of group identity (Ledgerwood and Liviatan, 2010; Mannetti et al., 2010). Out of these papers, those published in 2010 (except Conley et al., 2010; Echterhoff and Higgins, 2010) are contributions to a special issue on shared reality in the journal *Social Cognition*, for which I served as guest editor (see Echterhoff, 2010).

Testifying to the high political and societal relevance of related approaches, Kashima and colleagues have revealed the role of shared reality in the transmission of cultural stereotypes about various social categories (e.g., men, women, football players) through [p. 193 ↓] interpersonal communication (for a review, see Kashima, 2008). For instance, Lyons and Kashima (2003) experimentally manipulated the presence and absence of shared reality about a cultural stereotype to examine the conditions under which the stereotype is transmitted in a serial reproduction chain, which is set up akin to a Chinese whispers game. Stereotype-consistent (versus stereotype-inconsistent) information about a social group was transmitted to a greater extent along the chain when participants were led to believe that their audience shared the stereotypical view

of the Jamayans that they had learned. In contrast, no such biased transmission of stereotype-consistent information was found when people believed that their audience did not share the stereotypical view of the Jamayans.

The tendency to create a shared reality with one's interaction partners can also reduce people's stereotypes and thus can have beneficial consequences, as suggested by work by Sinclair, Hardin and colleagues (Sinclair et al., 2005a, 2005b). These researchers found that participants' endorsement of stereotypes of African-Americans, including self-stereotypes, can shift toward the egalitarian (i.e., nonprejudiced) views of an interaction partner, spontaneously achieving an interpersonal shared reality. The studies also elucidated the role of affiliative motivation. Affiliative motivation varied as a function of either a situational induction (the partner's manipulated similarity or likeability) or participants' existing affiliative motivation (assessed by personality scales). Participants' views shifted more toward their interaction partner's ostensible views when the social relationship motive was strong (versus weak).

Regarding the social and motivational dynamics in small groups, Mannetti et al. (2010) demonstrated the role of shared reality as a resource that allows people to ward off threats to a positive social identity. According to their main hypothesis, evaluations of defectors; that is, people who leave their current group to join another group, vary inversely with the threat they pose to other members' sense of shared reality. Consistent with this hypothesis, Mannetti et al. found that group members who have a stronger sense of shared reality within the group experience the defection of a group member as less threatening.

Furthermore, studies using the established saying-is-believing paradigm have applied shared reality theory to applied fields and thus enhanced the ecological validity of the evidence. The research on shared reality in the saying-is-believing paradigm is particularly relevant to the applied domain of eyewitness retellings and memory. In studies by Kopietz et al. (2009) and Hellmann et al. (in press), student participants tuned their retelling of a witnessed incident to their audience's evaluation of the suspects in the incident. In the study by Kopietz et al. (2009) participants' own memories and judgments regarding the incident were more biased toward their audience when they were more motivated to create a shared view with a particular audience (a student with a similar versus a dissimilar academic background).

Furthermore, Hellmann et al. (2010, Exp. 2) found that the correlation between message (“saying”) and recall (“believing”) was significantly higher when participants’ experienced a high (versus low) degree of shared reality with their audience.

Applying shared reality theory to the workplace and organizational behavior, Echterhoff et al. (2009b) examined the role of other audience characteristics in the context of personnel assessment. Student communicators described an employee to either an equal-status audience (a student temp) or a higher status audience (a company board member). The higher status audience clearly possessed higher domain-specific expertise, such as professional competence in the assessment of employees. Although audience tuning occurred in both audience-status conditions, the memory bias from audience tuning was found only in the equal-status condition. Apparently, communicators were more willing to share reality with the equal-status audience than with the higher status audience. [p. 194 ↓] An extended measure of trust in the audience, which include epistemic components (e.g., trust in the audience's judgments in general and about other people in particular) and relational components (e.g., readiness to affiliate and be close), was also higher in the equal-status condition and statistically mediated the audience-status effect on memory bias.

These findings show that an audience's domain-specific expertise or status is not sufficient to motivate communicators to create a shared reality with the audience. Rather, the audience's epistemic and relational trustworthiness is more critical. The feelings of general trust and the readiness to connect and affiliate covered by the extended trust measure in Echterhoff et al. (2009b) cannot be reduced to mere expertise. What matters is whether communicators want to make an epistemic and relational connection to the audience.

Regarding its broader utility, shared reality theory draws attention to a potentially important everyday mechanism underlying the construction of culturally shared memories and knowledge – a basic mechanism for constructing social, cultural, and political beliefs (see Hausmann et al., 2008; Jost et al., 2007). Consider community members who trust one another, want to maintain relationships with one another, and are thus prepared and motivated to create a shared reality. When community member A is aware of community member B's view (presumable or actual) regarding some topic (e.g., his or her beliefs or attitudes about something), audience tuning during

interpersonal communication is likely to occur. Given that the audience tuning serves a shared reality goal, it will shape communicator A's own later memories and beliefs about the topic in the direction of the audience. Rather than remembering the topic information as originally received, communicator A will remember this information as represented in her or his tuned message. A similar process may then occur when another community member C, being aware of A's communicatively-formed belief, talks about the same topic to member A, and then community member D talks to C, and so on. As a result of continued communication on the topic, the community members will come to hold increasingly shared beliefs.

This process may occur not only for individuals as topic targets but also for groups as topic targets (see Hausmann et al., 2008; Klein et al., 2008; Lyons and Kashima, 2003), and this could be an important factor in the development of shared stereotypic beliefs about other groups. Creating a shared reality with communication partners, which is ubiquitous in everyday life, can thus create a shared but biased perspective on the world within a community.

Acknowledgments

The preparation of this chapter was supported by a grant from the German Research Foundation (Deutsche Forschungsgemeinschaft, reference number EC 317/2). I thank Jens Hellmann, Natlija Keck, and Cécile Schain for their help with preparing the chapter for publication.

References

- Adaval, R. and Wyer, R.S., Jr. Communicating about a social interaction: Effects on memory for protagonists' statements and nonverbal behaviors . *Journal of Experimental Social Psychology* , 40, 450–465. (2004) <http://dx.doi.org/10.1016/j.jesp.2003.08.001>
- Andersen, S.M. and Chen, S. The relational self: An interpersonal social-cognitive theory . *Psychological Review* , 109, 619–645. (2002) <http://dx.doi.org/10.1037/0033-295X.109.4.619>

Asch, S.E. (1951) Effects of group pressure upon the modification and distortion of judgments . In H.S. Guetzkow (ed.), Groups, Leadership and Men , pp. 177–190. Pittsburgh, PA: Carnegie Press.

Asch, S.E. (1952) Social Psychology . Englewood Cliffs, NJ: Prentice Hall. <http://dx.doi.org/10.1037/10025-000>

Austin, J.L. (1962) How to Do Things with Words . Oxford: Oxford University Press.

Bar-Tal, D. (1990) Group Beliefs: A Conception for Analyzing Group Structure, Processes, and Behavior . New York: Springer-Verlag. <http://dx.doi.org/10.1007/978-1-4612-3298-8>

Bar-Tal, D. (2000). Shared Beliefs in a Society: Social Psychological Analysis . Thousand Oaks, CA: Sage. <http://dx.doi.org/10.4135/9781452204697>

Baron-Cohen, S. (1995) Mindblindness: An Essay on Autism and Theory of Mind . Cambridge, MA: MIT Press.

Bartlett, F.C. (1932) Remembering: A Study in Experimental and Social Psychology . Cambridge: Cambridge University Press.

Baumeister, R.F. and Leary, M.R. The need to belong: Desire for interpersonal attachments as a fundamental human motivation . Psychological Bulletin , 117, 497–529. (1995) <http://dx.doi.org/10.1037/0033-2909.117.3.497>

Bem, D.J. Self perception: An alternative interpretation of cognitive dissonance phenomena . Psychological Review , 74, 183–200. (1967) <http://dx.doi.org/10.1037/h0024835>

Berger, P.L. and Luckmann, T. (1966) The Social Construction of Reality: A Treatise in the Sociology of Knowledge . Garden City, NY: Doubleday.

Berlyne, D.E. Uncertainty and epistemic curiosity . British Journal of Psychology , 53, 27–34. (1962) <http://dx.doi.org/10.1111/j.2044-8295.1962.tb00811.x>

Bless, H., Strack, F. and Walther, E. (2001) Memory as a target of social influence? Memory distortions as a function of social influence and metacognitive knowledge . In J.P. Forgas, ed. and K.D. Williams (eds), Social Influence: Direct and Indirect processes , pp. 167–183. Philadelphia: Psychology Press.

Bohman, J.F., ed. and Rehg, W. (eds) (1997) Deliberative Democracy . Cambridge, MA: MIT Press.

Brentano, F. (1974) Psychology from an Empirical Standpoint (Trans. A. Rancurrello, ed. , D. Terrell, ed. and L. McAlister, ed.). London: Routledge and Kegan. (Originally published 1874).

Brickman, P. (1978) Is it real? In J.H. Harvey, ed. , W. Ickes, ed. and R.F. Kidd (eds), New Directions in Attribution Research , 2, 5–34. Hillsdale, NJ: Erlbaum.

Byrne, D. and Clore, G.L. Effectance arousal and attraction . Journal of Personality and Social Psychology , 6, 1–18. (1967) <http://dx.doi.org/10.1037/h0024829>

Caporael, L.R. (2007) Evolutionary theory for social and cultural psychology . In A.W. Kruglanski, ed. and E.T. Higgins (eds), Social Psychology: Handbook of Basic Principles , 2nd Edition, pp. 3–18. New York: Guilford Press.

Caporael, L.R. (2010) Revolutionary Darwinism: Sociality is the ground . In G. Semin, ed. and G. Echterhoff (eds), Grounding Sociality: Neurons, Minds, and Culture , pp. 237–259. New York: Psychology Press.

Chiu, C.Y., Krauss, R.M. and Lau, I.Y.M. (1998) Some cognitive consequences of communication . In S.R. Fussell, ed. and R. Kreuz (eds), Social and Cognitive Approaches to Interpersonal Communication , pp. 259–278. Hillsdale, NJ: Erlbaum.

Clark, H.H. (1996) Using Language . New York: Cambridge University Press.

Clark, H.H. (2003) Pointing and placing . In S. Kita (ed.), Pointing: Where Language, Culture, and Cognition meet , pp. 243–268. Hillsdale, NJ: Erlbaum.

Clark, H.H. and Marshall, C.E. (1981) Definite reference and mutual knowledge . In A.K. Joshi, ed. , B. Webber, ed. , and I. Sag (eds), Elements of Discourse Understanding , pp. 10–63. Cambridge: Cambridge University Press.

Conley, T.D., Rabinowitz, J.L. and Hardin, C.D. O.J. Simpson as shared (and unshared) reality: The impact of consensually shared beliefs on interpersonal perceptions and task performance in different-and same-ethnicity dyads . Journal of Personality and Social Psychology , 99, 452–466. (2010) <http://dx.doi.org/10.1037/a0019274>

Cooley, C.H. (1964) Human Nature and the Social Order . New York: Schocken Books. (Originally published 1902.)

Cuc, A., Koppel, J. and Hirst, W. Silence is not golden: A case for socially shared retrieval-induced forgetting . Psychological Science , 18, 727–733. (2007) <http://dx.doi.org/10.1111/j.1467-9280.2007.01967.x>

Deutsch, M. and Gerard, H.B. A study of normative and informational social influences upon individual judgment . Journal of Abnormal and Social Psychology , 51, 629–636. (1955) <http://dx.doi.org/10.1037/h0046408>

de Waal, F. Putting the altruism back into altruism: The evolution of empathy . Annual Review of Psychology , 59, 279–300. (2008) <http://dx.doi.org/10.1146/annurev.psych.59.103006.093625>

Diener, E. and Seligman, M.E.P. Very happy people . Psychological Science , 13, 81–84. (2002) <http://dx.doi.org/10.1111/1467-9280.00415>

Dovidio, J.F., Gaertner, S.L., Kawakami, K. and Hodson, G. Why can't we just get along? Interpersonal biases and interracial distrust . Cultural Diversity and Ethnic Minority Psychology , 8, 88–102. (2002) <http://dx.doi.org/10.1037/1099-9809.8.2.88>

Dunbar, R.I.M. and Shultz, S. Evolution in the social brain . Science , 317, 1344–1347. (2007) <http://dx.doi.org/10.1126/science.1145463>

Echterhoff, G. Shared reality: Antecedents, processes, and consequences . Social Cognition , 28, 273–276. (2010) <http://dx.doi.org/10.1521/soco.2010.28.3.273>

Echterhoff, G., Groll, S. and Hirst, W. Tainted truth: Overcorrection for misinformation influence on eyewitness memory . *Social Cognition* , 25, 367–409. (2007) <http://dx.doi.org/10.1521/soco.2007.25.3.367>

Echterhoff, G. and Higgins, E.T. (2010) How communication shapes memory: Shared reality and implications for culture . In G.R. Semin, ed. and G. Echterhoff (eds), *Grounding Sociality: Neurons, Minds, and Culture* , pp. 115–146. New York: Psychology Press.

Echterhoff, G., Higgins, E.T. and Groll, S. Audience-tuning effects on memory: The role of shared reality . *Journal of Personality and Social Psychology* , 89, 257–276. (2005) <http://dx.doi.org/10.1037/0022-3514.89.3.257>

Echterhoff, G., Higgins, E.T., Kopietz, R. and Groll, S. How communication goals determine when audience tuning biases memory . *Journal of Experimental Psychology: General* , 137, 3–21. (2008) <http://dx.doi.org/10.1037/0096-3445.137.1.3>

Echterhoff, G., Higgins, E.T. and Levine, J.M. Shared reality: Experiencing commonality with others' inner states about the world . *Perspectives on Psychological Science* , 4, 496–521. (2009a) <http://dx.doi.org/10.1111/j.1745-6924.2009.01161.x>

Echterhoff, G. and Hirst, W. Social influence on memory . *Social Psychology* , 40, 106–110. (2009) <http://dx.doi.org/10.1027/1864-9335.40.3.106>

Echterhoff, G., Lang, S., Krämer, N. and Higgins, E.T. Audience-tuning effects on communicators' memory: The role of audience status in sharing reality . *Social Psychology* , 40, 150–163. (2009b) <http://dx.doi.org/10.1027/1864-9335.40.3.150>

Epley, N., Caruso, E.M. and Bazerman, M.H. When perspective-taking increases taking: Reactive egoism in social interaction . *Journal of Personality and Social Psychology* , 91, 872–889. (2006) <http://dx.doi.org/10.1037/0022-3514.91.5.872>

Festinger, L. Informal social communication . *Psychological Review* , 57, 271–282. (1950) <http://dx.doi.org/10.1037/h0056932>

Festinger, L. (1957) *A Theory of Cognitive Dissonance* . Stanford: Stanford University Press.

Fiske, S.T. (2007) Core social motivations, a historical perspective: Views from the couch, consciousness, classroom, computers, and collectives . In W. Gardner, ed. and J.Y. Shah (eds), *Handbook of Motivation Science* , pp. 3–22. New York: Guilford Press.

Frith, C.D. (2007) *Making up the Mind: How the Brain Creates Our Mental World* . Malden, MY: Blackwell.

Gallese, V., Keysers, C. and Rizzolatti, G. A unifying view of the basis of social cognition . *Trends in Cognitive Sciences* , 8, 396–403. (2004) <http://dx.doi.org/10.1016/j.tics.2004.07.002>

Garfinkel, H. (1967) *Studies in Ethnomethodology* . Englewood Cliffs, NJ: Prentice-Hall.

Graf, M., Schütz-Bosbach, S. and Prinz, W. (2010) Motor involvement in action and object perception: Similarity and complementarity . In G. Semin, ed. and G. Echterhoff (eds), *Grounding Sociality: Neurons, Minds, and Culture* , pp. 27–52. New York: Psychology Press.

Grice, H.P. (1975) Logic and conversation . In P. Cole, ed. and J.L. Morgan (eds), *Syntax and Semantics 3: Speech Acts* , pp. 41–58. San Diego: Academic Press.

Hardin, C.D. and Conley, T.D. (2001) A relational approach to cognition: Shared experience and relationship affirmation in social cognition . In G.B. Moskowitz (ed.), *Cognitive Social Psychology: The Princeton Symposium on the Legacy and Future of Social Cognition* , pp. 3–17. Mahwah, NJ: Erlbaum.

Hardin, C.D. and Higgins, E.T. (1996) Shared reality: How social verification makes the subjective objective . In R.M. Sorrentino, ed. and E.T. Higgins (eds), *Handbook of Motivation and Cognition: The Interpersonal Context* , 3, 28–84. New York: Guilford Press.

Hausmann, L.R.M., Levine, J.M. and Higgins, E.T. Communication and group perception: Extending the ‘saying is believing’ effect . *Group Processes and Intergroup Relations* , 11, 539–554. (2008) <http://dx.doi.org/10.1177/1368430208095405>

Heider, F. (1958) *The Psychology of Interpersonal Relations* . New York: John Wiley.
<http://dx.doi.org/10.1037/10628-000>

Hellmann, J.H., Echterhoff, G., Kopietz, R., Niemeier, S. and Memon, A. (in press) Talking about visually perceived events: Communication effects on eyewitness memory . *European Journal of Social Psychology* .

Higgins, E.T. (1981) The ‘communication game’: Implications for social cognition and persuasion . In E.T. Higgins, ed. , C.P. Herman, ed. , and M.P. Zanna (eds), *Social Cognition: The Ontario Symposium* , 1, 343–392. Hillsdale, NJ: Erlbaum.

Higgins, E.T. Achieving ‘shared reality’ in the communication game: A social action that creates meaning . *Journal of Language and Social Psychology* , 11, 107–131. (1992)
<http://dx.doi.org/10.1177/0261927X92113001>

Higgins, E.T. (1996a) Knowledge activation: Accessibility, applicability, and salience . In E.T. Higgins, ed. and A.W. Kruglanski (eds), *Social Psychology: Handbook of Basic Principles* , pp. 133–168. New York: Guilford Press.

Higgins, E.T. (1996b) Shared reality in the self-system: The social nature of self-regulation . In W. Stroebe, ed. and M. Hewstone (eds), *European Review of Social Psychology* , 7, 1–29. Chichester: Wiley. <http://dx.doi.org/10.1080/14792779443000085>

Higgins, E.T. The aboutness principle: A pervasive influence on human inference . *Social Cognition* , 16, 173–198. (1998) <http://dx.doi.org/10.1521/soco.1998.16.1.173>

Higgins, E.T. (1999) ‘Saying is believing’ effects: When sharing reality about something biases knowledge and evaluations . In L.L. Thompson, ed. , J.M. Levine, ed. and D.M. Messick (eds), *Shared Cognition in Organizations: The Management of Knowledge* , pp. 33–49. Mahwah, NJ: Erlbaum.

Higgins, E.T. (2005) Humans as applied motivation scientists: Self-consciousness from 'shared reality' and 'becoming' . In H.S. Terrace, ed. and J. Metcalfe (eds), *The Missing Link in Cognition: Origins of Self-reflective Consciousness* , pp. 157–173. Oxford: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780195161564.003.0006>

Higgins, E.T. Value from hedonic experience and engagement . *Psychological Review* , 113, 439–460. (2006) <http://dx.doi.org/10.1037/0033-295X.113.3.439>

Higgins, E.T. (in press) Beyond Pleasure and Pain: The New Science of Motivation . New York: Oxford University Press.

Higgins, E.T. (2010) Sharing inner states: A defining feature of human motivation . In G. Semin, ed. and G. Echterhoff (eds), *Grounding Sociality: Neurons, Minds, and Culture* , pp. 149–173. New York: Psychology Press.

Higgins, E.T., Echterhoff, G., Crespillo, R. and Kopietz, R. Effects of communication on social knowledge: Sharing reality with individual versus group audiences . *Japanese Psychological Research* , 49, 89–99. (2007) <http://dx.doi.org/10.1111/j.1468-5884.2007.00336.x>

Higgins, E.T. and McCann, C.D. Social encoding and subsequent attitudes, impressions, and memory: 'Context-driven' and motivational aspects of processing . *Journal of Personality and Social Psychology* , 47, 26–39. (1984) <http://dx.doi.org/10.1037/0022-3514.47.1.26>

Higgins, E.T., McCann, C.D. and Fondacaro, R. The 'communication game': Goal directed encoding and cognitive consequences . *Social Cognition* , 1, 21–37. (1982) <http://dx.doi.org/10.1521/soco.1982.1.1.21>

Higgins, E.T. and Pittman, T.S. Motives of the human animal: Comprehending, managing, and sharing inner states . *Annual Review of Psychology* , 59, 361–385. (2008) <http://dx.doi.org/10.1146/annurev.psych.59.103006.093726>

Higgins, E.T. and Rholes, W.S. 'Saying is believing': Effects of message modification on memory and liking for the person described . *Journal*

of Experimental Social Psychology , 14, 363–378. (1978) <http://dx.doi.org/10.1016/0022-1031%2878%2990032-X>

Hirst, W. and Echterhoff, G. Creating shared memories in conversation: Toward a psychology of collective memory . Social Research , 75, 183–216. (2008)

Hirst, W. and Manier, D. (2002) The diverse forms of collective memory . In G. Echterhoff, ed. and M. Saar (eds), Kontexte und Kulturen des Erinnerns [Contexts and cultures of remembering] , pp. 37–58. Constance: Universitätsverlag Konstanz.

Hogg, M.A. (2007) Uncertainty-identity theory . In M.P. Zanna (ed.), Advances in Experimental Social Psychology , 39, 69–126. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2806%2939002-8>

Holtgraves, T.M. and Kashima, Y. Language, meaning, and social cognition . Personality and Social Psychology Review , 12, 73–94. (2008) <http://dx.doi.org/10.1177/1088868307309605>

Hunt, E. and Agnoli, F. The Whorfian hypothesis: A cognitive psychology perspective . Psychological Review , 98, 377–389. (1991) <http://dx.doi.org/10.1037/0033-295X.98.3.377>

Huntsinger, J.R. and Sinclair, S. When it feels right, go with it: Affective regulation of affiliative social tuning . Social Cognition , 28, 287–302. (2010) <http://dx.doi.org/10.1521/soco.2010.28.3.290>

Husserl, E. (1931) Ideas: General Introduction to Pure Phenomenology (Trans. W.R.B. Gibson, ed.). London: George Allen & Unwin. (Originally published 1913.)

Iacoboni, M. (2008) Mirroring People: The New Science of How We Connect with Others . New York: Farrar, Straus, and Giroux.

Iacoboni, M. (in press) Mirroring as a key neural mechanism of sociality . In G. Semin, ed. and G. Echterhoff (eds), Grounding Sociality: Neurons, Minds, and Culture . New York: Psychology Press.

James, W. (1890) *The Principles of Psychology*, Vol. 1 . Cambridge, MA: Harvard University Press. <http://dx.doi.org/10.1037/10538-000>

Johnson, M.K. and Raye, C.L. Reality monitoring . *Psychological Review* , 88, 67–85. (1981) <http://dx.doi.org/10.1037/0033-295X.88.1.67>

Jones, E.E. (1964) *Ingratiation: A Social Psychological Analysis* . New York: Appleton-Century-Crofts.

Jones, E.E. and Thibaut, J.W. (1958) Interaction goals as bases of inference in interpersonal perception . In R. Tagiuri, ed. and L. Petrullo (eds), *Person Perception and Interpersonal Behavior* , pp. 151–178. Stanford: Stanford University Press.

Jost, J.T., Ledgerwood, A. and Hardin, C.D. Shared reality, system verification, and the relational basis of ideological beliefs . *Social and Personality Psychology Compass* , 2, 171–186. (2007) <http://dx.doi.org/10.1111/j.1751-9004.2007.00056.x>

Kagan, J. Motives and development . *Journal of Personality and Social Psychology* , 22, 51–66. (1972) <http://dx.doi.org/10.1037/h0032356>

Kashima, Y. A social psychology of cultural dynamics: Examining how cultures are formed, maintained, and transformed . *Social and Personality Psychology Compass* , 2, 107–120. (2008) <http://dx.doi.org/10.1111/j.1751-9004.2007.00063.x>

Kashima, Y., KashimaE.S., Bain, P., Lyons, A., Tindale, R.S., Robins, G., Vears, C. and Whelan, J. Com munication and essentialism: Grounding the shared reality of a social category . *Social Cognition* , 28, 306–328. (2010) <http://dx.doi.org/10.1521/soco.2010.28.3.306>

Keysar, B. and Barr, D.J. (2002) Self anchoring in conversation: Why language users do not do what they ‘should’ . In T. Gilovich, ed. , D.W. Griffin, ed. , and D. Kahneman (eds), *Heuristics and Biases: The Psychology of Intuitive Judgment* , pp. 150–166. Cambridge: Cambridge University Press.

Keysers, C. and Gazzola, V. Integrating simulation and theory of mind: From self to social cognition . Trends in Cognitive Sciences , 11, 194–196. (2007) <http://dx.doi.org/10.1016/j.tics.2007.02.002>

Klein, O., Clark, A.E. and Lyons, A. When the social becomes personal: Exploring the role of common ground in stereotype communication . Social Cognition , 28, 329–352. (2010) <http://dx.doi.org/10.1521/soco.2010.28.3.329>

Klein, O., Tindale, S. and Brauer, M. (2008) The consensualization of stereotypes in small groups . In Y. Kashima, ed. , K. Fiedler, ed. , and P. Freytag (eds), Stereotype Dynamics: Language-Based Approaches to the Formation, Maintenance, and Transformation of Stereotypes , pp. 263–292. New York: Erlbaum.

Kopietz, R., Echterhoff, G., Niemeier, S., Hellmann, J.H. and Memon, A. Audience-congruent biases in eyewitness memory and judgment: Influences of a co-witness' liking for a suspect . Social Psychology , 40, 138–149. (2009) <http://dx.doi.org/10.1027/1864-9335.40.3.138>

Kopietz, R., Hellmann, J.H., Higgins, E.T. and Echterhoff, G. Shared reality effects on memory: Communicating to fulfill epistemic needs . Social Cognition , 28, 353–378. (2010) <http://dx.doi.org/10.1521/soco.2010.28.3.353>

Kruglanski, A.W., Pierro, A., Mannetti, L. and De Grada, E. Groups as epistemic providers: Need for closure and the unfolding of group-centrism . Psychological Review , 113, 84–100. (2006) <http://dx.doi.org/10.1037/0033-295X.113.1.84>

Latané, B. and Wolf, S. The social impact of majorities and minorities . Psychological Review , 5, 438–453. (1981)

Ledgerwood, A. and Liviatan, I. The price of a shared vision: Group identity as goals and the social creation of value . Social Cognition , 28, 401–421. (2010) <http://dx.doi.org/10.1521/soco.2010.28.3.401>

Levine, J.M. and Higgins, E.T. (2001) Shared reality and social influence in groups and organizations . In F. Butera, ed. and G. Mugny (eds), Social Influence in Social Reality: Promoting Individual and Social Change , pp. 33–52. Seattle: Hogrefe & Huber.

Lewin, K. (1947) Group decision and social change . In T.M. Newcomb, ed. and E.L. Hartley (eds), *Readings in Social Psychology* , pp. 330–344. New York: Holt.

Lindner, I., Echterhoff, G., Davidson, P.S.R. and Brand, M. Observation inflation: Your actions become mine . *Psychological Science* , 21, 1291–1299. (2010) <http://dx.doi.org/10.1177/0956797610379860>

Loewenstein, G. The psychology of curiosity: A review and reinterpretation . *Psychological Bulletin* , 116, 75–98. (1994) <http://dx.doi.org/10.1037/0033-2909.116.1.75>

Lyons, A. and Kashima, Y. How are stereotypes maintained through communication? The influence of stereotype sharedness . *Journal of Personality and Social Psychology* , 85, 989–1005. (2003) <http://dx.doi.org/10.1037/0022-3514.85.6.989>

Mackie, D.M. and Skelly, J.J. (1994) The social cognition analysis of social influence: Contributions to the understanding of persuasion and conformity . In P. Devine, ed. , D. Hamilton, ed. , and T. Ostrom (eds), *Social Cognition: Impact on Social Psychology* , pp. 259–289. New York: Academic Press.

Magee, M.W. and Hardin, C.D. In defense of religion: Shared reality moderates the unconscious threat of evolution . *Social Cognition* , 28, 379–400. (2010) <http://dx.doi.org/10.1521/soco.2010.28.3.379>

Malle, B.F. How people explain behavior: A new theoretical framework . *Personality and Social Psychology Review* , 3, 21–43. (1999) http://dx.doi.org/10.1207/s15327957pspr0301_2

Malle, B.F., ed. and Hodges, S.D. (eds) (2005) *Other Minds: How Humans Bridge the Divide Between Self and Other* . New York: Guilford Press.

Mannetti, L., Levine, J.M., Pierro, A. and Kruglanski, A.W. Group reaction to defection: The impact of shared reality . *Social Cognition* , 28, 447–464. (2010) <http://dx.doi.org/10.1521/soco.2010.28.3.447>

Marsh, E.J. Retelling is not the same as recalling: Implications for memory . Current Directions in Psychological Science , 16, 16–20. (2007) <http://dx.doi.org/10.1111/j.1467-8721.2007.00467.x>

Marsh, K.L. (2010) Sociality, from an ecological, dynamical perspective . In G. Semin, ed. and G. Echterhoff (eds), Grounding Sociality: Neurons, Minds, and Culture . pp. 53–81. New York: Psychology Press.

McCann, C.D. and Higgins, E.T. (1992) Personal and contextual factors in communication: A review of the ‘communication game’ . In G.R. Semin, ed. and K. Fiedler (eds), Language, Interaction and Social Cognition , pp. 144–171. London: Sage.

Mead, G.H. (1934) Mind, Self, and Society . Chicago: University of Chicago Press.

Meltzoff, A.N. and Decety, J. What imitation tells us about social cognition: A rapprochement between developmental psychology and cognitive neuroscience . Philosophical Transactions of the Royal Society London, Series B , 358, 491–500. (2003) <http://dx.doi.org/10.1098/rstb.2002.1261>

Merton, R.K. and Kitt, A. (1950) Contributions to the theory of reference group behavior . In R.K. Merton, ed. and P.F. Lazarsfeld (eds), Continuities in Social Research: Studies in the Scope and Method of ‘The American Soldier’ , pp. 40–105. Glencoe, IL: Free Press.

Moscovici, S. (1981) On social representations . In J.P. Forgas (ed.), Social Cognition: Perspectives on Everyday Understanding , pp. 181–209. London: Academic Press.

Neumann, R. and Strack, F. ‘Mood contagion’: The automatic transfer of mood between persons . Journal of Personality and Social Psychology , 79, 211–223. (2000) <http://dx.doi.org/10.1037/0022-3514.79.2.211>

Newcomb, T.M. (1959) Individual systems of orientation . In S. Koch (ed.), Psychology: A Study of a Science: Formulations of the Person and the Social Context , 3, 384–422. New York: McGraw-Hill.

Nickerson, R.S. The projective way of knowing: A useful heuristic that sometimes misleads . Current Directions in Psychological Science , 10, 168–172. (2001) <http://dx.doi.org/10.1111/1467-8721.00141>

Ostrom, T.M. (1984) The sovereignty of social cognition . In R.S. Wyer, Jr., ed. and T.K. Srull (eds), Handbook of Social Cognition , 1, 1–38. Hillsdale, NJ: Erlbaum.

Pickering, M.J. and Garrod, S. Toward a mechanistic psychology of dialogue . Behavioral and Brain Sciences , 27, 169–225. (2004)

Pinel, E.C., Long, A.E. and Crimin, L.A. I-sharing and a classic conformity paradigm . Social Cognition , 28, 277–289. (2010) <http://dx.doi.org/10.1521/soco.2010.28.3.277>

Resnick, L.B., ed. , Levine, J.M., ed. and Teasley, S.D. (eds) (1991) Perspectives on Socially Shared Cognition . Washington, DC: American Psychological Association.
<http://dx.doi.org/10.1037/10096-000>

Richeson, J.A. and Trawalter, S. Why do interracial interactions impair executive function? A resource depletion account . Journal of Personality and Social Psychology , 88, 934–947. (2005) <http://dx.doi.org/10.1037/0022-3514.88.6.934>

Rommetveit, R. (1974) On Message Structure: A Framework for the Study of Language and Communication . New York: Wiley.

Salas, E., ed. and Fiore, S.M. (eds) (2004) Team Cognition: Understanding the Factors That Drive Process and Performance . Washington, DC: American Psychological Association.

Sapir, E. (1964) Culture, Language, and Personality . Berkeley, CA: University of California Press. (Originally published 1941.)

Schachter, S. (1959) The Psychology of Affiliation: Experimental Studies of the Sources of Gregariousness . Stanford, CA: Stanford University Press.

Schütz, A. (1967) The Phenomenology of the Social World . Evanston, IL: Northwestern University Press. (Originally published 1932.)

Sedikides, C. Efforts of fortuitously activated constructs versus activated communication goals on person impressions . *Journal of Personality and Social Psychology* , 58, 397–408. (1990) <http://dx.doi.org/10.1037/0022-3514.58.3.397>

Semin, G.R. (2007) Grounding communication: Synchrony . In A.W. Kruglanski, ed. and E.T. Higgins (eds), *Social Psychology: Handbook of Basic Principles* , 2nd Edition, pp. 630–649. New York: Guilford Press.

Sherif, M. A study of some social factors in perception . *Archives of Psychology* , 27 (Whole No. 187) (1935)

Sherif, M. (1936) *The Psychology of Social Norms* . New York: Harper and Brothers.

Sinclair, S., Huntsinger, J., Skorinko, J. and Hardin, C.D. Social tuning of the self: Consequences for the self-evaluations of stereotype targets . *Journal of Personality and Social Psychology* , 89, 160–175. (2005a) <http://dx.doi.org/10.1037/0022-3514.89.2.160>

Sinclair, S., Lowery, B.S., Hardin, C.D. and Colangelo, A. Social tuning of automatic racial attitudes: The role of affiliative motivation . *Journal of Personality and Social Psychology* , 89, 583–592. (2005b) <http://dx.doi.org/10.1037/0022-3514.89.4.583>

Smith, E.R. and Semin, G.R. (2004) Socially situated cognition: Cognition in its social context . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 36, 53–115. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2804%2936002-8>

Stukas, A., Bratanova, B., Kashima, Y., Peters, K. and Beatson, R. The effect of social tuning and shared reality on recipients' beliefs and impressions . *Social Influence* , 5, 101–117. (2010) <http://dx.doi.org/10.1080/15534510903384916>

Sullivan, H.S. (1953) *The Interpersonal Theory of Psychiatry* . New York: Norton.

Thagard, P. Collaborative knowledge . *Nous* , 31, 242–261. (1997) <http://dx.doi.org/10.1111/0029-4624.00044>

Thompson, L. and Fine, G.A. Socially shared cognition, affect, and behavior: A review and integration . *Personality and Social Psychology Review* , 3, 278–302. (1999) http://dx.doi.org/10.1207/s15327957pspr0304_1

Tomasello, M. (2008) *Origins of Human Communication* . Cambridge, MA: MIT Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T. and Moll, H. Understanding and sharing intentions: The origins of cultural cognition . *Behavioral and Brain Sciences* , 28, 675–691. (2005)

Turner, J.H. Toward a sociological theory of motivation . *Annual Sociological Review* , 52, 15–27. (1987) <http://dx.doi.org/10.2307/2095389>

Tversky, B. and Marsh, E.J. Biased retellings of events yield biased memories . *Cognitive Psychology* , 40, 1–38. (2000) <http://dx.doi.org/10.1006/cogp.1999.0720>

Vedantam, S. (2008) For political candidates, saying can become believing . *The Washington Post* , February 25, A03.

Weldon, M.S. (2001) Remembering as a social process . In D.L. Medin (ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* , 40, 67–120. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0079-7421%2800%2980018-3>

Whorf, B.L. (1956) *Language Thought, and Reality: Selected Writings of Benjamin Lee Whorf* (John B. Carroll, ed.). Cambridge, MA: MIT Press.

Wittgenstein, L. (1953) *Philosophical Investigations* . New York: Macmillan.

<http://dx.doi.org/10.4135/9781446249222.n35>

Handbook of Theories of Social Psychology

Equity Theory in Close Relationships

Contributors: Richard L. Rapson & Elaine Hatfield

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Equity Theory in Close Relationships"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n36>

Print pages: 200-218

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n36>

[p. 200 ↓]

Chapter 36: Equity Theory in Close Relationships

Richard L.Rapson ElaineHatfield

Abstract

Throughout history, people have been concerned with social justice. In the eleventh century, St. Anselm of Canterbury (1998) argued that the will possesses two competing inclinations: an affection for a person's own advantage *and* an affection for justice. The first inclination is stronger, but the second matters, too. Equity theory, too, posits that in personal relationships, two concerns stand out: First, how rewarding are one's societal, family, and work relationships? Second, how fair and equitable are those relationships? According to equity theory, people feel most comfortable when they are getting exactly what they deserve from their relationships – no more and certainly no less. In this paper, we begin by describing the social concerns that sparked our interest in developing a theory of social justice. Then we describe the classic equity paradigm and the research it fostered. We recount the great intellectual debate that arose in the wake of the assertion that even in close, loving relationships, both reward and fairness matter. We end by describing current multicultural and multidisciplinary research that lends a new richness to theories of social justice, and contributes to the theory's usefulness in addressing current social issues.

Introduction

In the West, the 1960s and 1970s were a time of intellectual and social ferment. There was a great concern with social justice and spirited debate as to what was fair in life, law, marriage, and work. In the US, it was the time of Martin Luther King's historic 1965

civil rights march from Selma to Montgomery. (On “Bloody Sunday,” March 7, 1965, 600 civil rights marchers were attacked by state and local police with clubs, dogs, and tear gas.) It was the time of Jane Fonda’s 1972 trip to North Vietnam to protest the war. In that same year, women lobbied, marched, petitioned, picketed, and committed acts of civil disobedience in the hopes of persuading the 92nd Congress to pass the Equal Rights Amendment, which guaranteed men and women equal rights under law. (It passed the Senate and the House, but in the end the states failed to ratify it.) It was an era when feminists such as Betty Friedan described the *Feminine Mystique*, Gloria Steinem and her colleagues founded *Ms. Magazine*, and Shulamith Firestone penned *The Dialectic of Sex*. All these feminist leaders argued for women’s rights in education, law, and the workplace. On the comic side, Bobby Riggs spewed out chauvinist insults in challenging tennis star Billie Jean King to the “Battle of the Sexes.” (King won handily.) Valerie Solanas contributed her mad ravings to the *SCUM Manifesto*. (SCUM = The Society for [p. 201 ↓] Cutting Up Men.) (We assumed Ms. Solanas was a witty satirist until she put her money where her mouth was and shot her pal Andy Warhol.)

Personal Narrative of Equity Theory's Development

Given the times, it is not surprising that many of our friends and colleagues began to hotly debate the nature of social justice and the role that reward and fairness play in men and women’s close intimate relationships. Opinions and experiences differed widely. A few older women admitted they were grateful that their husbands allowed them the luxury of working. Thus, they went out of their way to make sure his masculinity was never threatened by their professional commitments – to ensure that their husbands never came home to a tired, disheveled, or cross wife; that the men were never “stuck” with childcare, housework, or yard work. A few of my more spirited younger friends were sick and tired of such inequities and argued that women ought to demand a marriage contract to ensure marital fairness.

Intrigued by all this speculation, Elaine Hatfield, G. William Walster, and Ellen Berscheid (1978) set out to devise a theory as to what men and women perceived to be fair in their daily encounters and the consequences of such perceptions. Ideally, we hoped

to devise a theory that would be applicable to all cultures and all historical eras. We believed that a concern with fairness was a cultural universal. We were convinced that during humankind's long evolutionary heritage, a concern with social justice came to be writ in the mind's "architecture" because such values possessed survival value. Such concerns were maintained, we thought, because behaving fairly continued to be a wise and profitable strategy in today's world¹ (For a further discussion of this point, see Hatfield et al., 2008; Tooby and Cosmides, 1992).

Yet, we were also aware that, throughout history, societies have had very different visions as to what constitutes "social justice," "fairness," and "equity." Some dominant views:

- "All men are created equal" (equality).
- "The more you invest in a project, the more profit you deserve to reap" (contemporary American capitalism).
- "To each according to his need" (communism).
- "Winner take all" (dog-eat-dog capitalism).
- It's a man's world (traditional patriarchy). In fifteenth-century England, we knew, the status hierarchy was God, men, farm animals (especially horses), then women and children.

Thus, in crafting equity theory, we attempted to create a model that would allow scholars to take men and women's own social perspectives into account when defining reward and fairness and justice. We came up with this model ([Figure 36.1](#)).

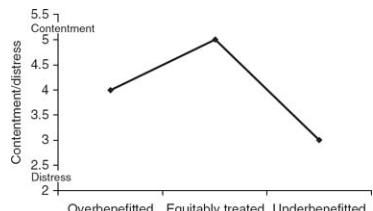
Equity Theory and Research

Equity theory is a straightforward theory. It consists of four propositions:

- *Proposition I:* Men and women are "hardwired" to try to maximize pleasure and minimize pain.
- *Proposition II:* Society, however, has a vested interest in persuading people to behave fairly and equitably. Groups will generally reward members who treat others equitably and punish those who treat others inequitably.

- *Proposition III:* Given societal pressures, people are most comfortable when they perceive that they are getting roughly what they deserve from life and love. If people feel over-benefited, they may experience pity, guilt, and shame.² If they feel under-benefited, they may experience anger, sadness, and resentment.
- *Proposition IV:* People in inequitable relationships will attempt to reduce their distress through a variety of techniques – by restoring psychological equity, actual equity, or leaving the relationship.

Figure 36.1 The relationship between equity and inequity, and dating and marital satisfaction



Assessing Equity

In practice, a relationship's fairness and equity can be reliably and validly assessed with the use of a simple measure – the Global Measure of Equity. Specifically, research participants are asked: “Considering what you put into your dating relationship or marriage, compared to what you get out of it ... and what your partner puts in compared to what (s)he gets out of it, how does your dating relationship or marriage ‘stack up’”? Respondents are given the following response options:

- +3: I am getting a much better deal than my partner.
- +2: I am getting a somewhat better deal.
- +1: I am getting a slightly better deal.
- 0: We are both getting an equally good, or bad, deal.
- -1: My partner is getting a slightly better deal.
- -2: My partner is getting a somewhat better deal.
- -3: My partner is getting a much better deal than I am.

On the basis of their answers, persons can be classified as over-benefited (receiving more than they deserve), equitably treated, or under-benefited (receiving less than they deserve).

Other, more detailed measures of equity exist, of course. Hatfield and her colleagues (2008) asked men and women who were dating, living together, and married to indicate how fair and equitable they considered their relationships to be via a 25-item scale – the Multi-Factor Measure of Equity. The areas of interest included such *personal concerns* as appearance, intelligence, and social grace; *emotional concerns*, such as physical affection and understanding and concern; and *day-to-day concerns*, such as contributing to household expenses and helping around the house. (See Young and Hatfield, 2009, for information on the reliability and validity of both these measures.)

Regardless of societal definitions or one's own concern with equity, considerations of equity have been found to be important in a wide variety of cultures and relationships – social relationships, romantic and family relationships, friendships, helping relationships, and work relationships.

The Importance of Equity in Close Intimate Relationships

Social psychologists are well aware that relationships change and deepen over time (Hatfield and Rapson, 1993). In *Intimate Relationships*, for example, Perlman and Duck (1986) argued that relationships go through stages. They charted the initiation, maintenance, problems, repair, and termination of relationships. Equity theorists, too, have been interested in charting the degree to which couples' concerns with reward, fairness, and equity wax and wane during the course of a love affair. Scholars have discovered that how concerned couples are with reward and equity depends on the stage at which their relationship has arrived. When couples are first dating, they participate in a kind of "dating and marriage marketplace," in which considerations of reward, fairness, and equity loom large. Once men and women are deeply committed, however, they become less concerned about day-to-day reward and equity. Should a relationship deteriorate, couples – knowing (perhaps) that they will soon be back on the

dating and marriage market – may begin to worry about, “What's in it for me?” and to ask, “Do I deserve better?”

Let us review the research leading to these conclusions.

Beginnings

In fairy tales, Prince Charming often falls in love with the scullery maid. In real life, however, dating couples – whether they are young or old, gay, lesbian, or heterosexual – generally search for “suitable” partners. As Goffman [p. 203 ↓] observed: “A proposal of marriage in our society tends to be a way in which a man sums up his social attributes and suggests that hers are not so much better as to preclude a merger” (1952: 456).

Specifically, researchers find:

- Attractive men and women assume that a suitable partner should be more socially desirable – more attractive, intelligent, personable, rich, well adjusted, and kind – than do their less attractive peers.
- Dating couples are more likely to fall in love if they perceive their relationships to be fair and equitable.
- Couples are likely to end up with someone fairly close to themselves in social desirability. They are likely to be matched on the basis of self-esteem, looks, intelligence, education, and mental and physical health (or disability). Evolutionary theorists contend that in the dating marketplace, men are willing to pay a price for good looks, virginity, fidelity, and chastity, while women willingly pay for status, support, and kindness. Market considerations have also been found to affect the amount prostitutes charge for risky sex, the sexual bargains men and women craft in prison, and the like.
- Perceived equity is important in sparking passionate love, sexual attraction, sexual passion, and sexual activity.
- Equitable dating relationships are satisfying and comfortable relationships; inequity is associated with distress, guilt, anger, and anxiety.
- Equitable dating relationships are more stable than are inequitable relationships.

In conclusion, research seems to indicate that in the *early stages* of a dating relationship, considerations of the marketplace prevail. Men and women attempt to attract socially attractive partners and are profoundly concerned with how rewarding, fair, and equitable their budding relationships appear to be. (Additional support for these propositions can be found in Baumeister and Vohs, 2004.)

Flowerings

In *Equity: Theory and Research*, Hatfield and her colleagues (1978) argued that casual dating relationships differ from deeply committed, loving, intimate relationships in several ways. Specifically:

- 1. intensity of liking and loving;
- 2. depth and breadth of information exchange;
- 3. length of relationship;
- 4. value of resources exchanged;
- 5. variety of resources exchanged;
- 6. interchangeability of resources;
- 7. the unit of analysis: from “you” and “me” to “we” (1978: 183).

Long married couples, who assume they will be together for a lifetime, are likely to be fairly sanguine about momentary injustices, confident that it will all work out in the end. Also, given the complexity of marital relationships, it may be difficult for married couples to calculate moment-to-moment whether or not their relationships are fair. They may well settle for a rough and ready definition of fairness. (“Yeah, all-in-all, things seem pretty fair to me, I guess.”) Love might also affect how people caught up in inequitable relationships go about trying to set momentary inequities right.

Yet, in the end – in even the closest of relationships – fairness and equity do matter. Most intimates assume that good deeds will eventually be rewarded. (Their partners will be grateful. They will love them more. They will wish to reciprocate.) When people are forced to suffer too much, for too long, with no hope of return, they may well begin to resent life's unfairness. The wife of the Alzheimer's patient may begin to ask “Why me?” and to wish she could be released from her terrible burden. Her husband may feel guilty

upon contemplating her terrible plight; he, too, may cry, shamefaced: "It's not fair" (Clark and Grote, 1998; Hatfield, et al., 1978, 2008; Markman, 1981).

Scientists have found that most couples – single, living together, or married; affluent or poor; dating for a few weeks or married for 50 years – do care about equity. In all of these groups, degree of reward, fairness, and equity have been found to be linked to passionate and companionate love, sexual satisfaction, marital happiness, contentment, [p. 204 ↓] satisfaction, and marital stability (Buunk and van Yperen, 1989; Byers and Wang, 2004; Hatfield et al., 2008; Lawrence and Byers, 1995; Martin, 1985; Schreurs and Buunk, 1996; van Yperen and Buunk, 1990). Couples in equitable relationships are also less likely to risk extramarital affairs than are their peers. They are more confident that their marriages will last, and in fact their relationships are longer lasting than those of their peers (Byers and Wang, 2004; Hatfield, et al., 2008; van Yperen and Buunk, 1990).

In one longitudinal study, Pillemer et al. (2008), interviewed a sample of dating and married men and women in Madison, Wisconsin, who ranged in age from 18 to 92; the couples been married between 1 and 53 years. They found that:

- Older women were hesitant to talk about fairness and equity in a marriage. They felt that couples shouldn't think in selfish ways – worrying about whether or not they were getting their fair share.
- The vast majority of older women (85 percent) considered their marriages to be fair and equitable.
- Older women appeared to be less concerned about day-to-day inequities than were dating couples and newlyweds. Nonetheless, even in the best of marriages, most admitted that niggling doubts about fairness did surface now and then. As predicted, those who were over-benefited felt a bit guilty; those who were under-benefited felt far more angry than did their privileged peers.
- Stressful life events – such as the arrival of children, retirement, serious illness, or the awareness of impending death – often brought to awareness long simmering resentments over issues of fairness.

In conclusion, research seems to indicate that although men and women who have been in a close intimate relationship for a long period of time are more tolerant of

inequity than are their peers, in the end, they appear to be deeply concerned with how rewarding and equitable their relationships are.

The End of the Affair

Hatfield and colleagues (1978) argued that if men and women are unfairly treated for a prolonged period of time, they will begin to wonder: "Does my partner love me? If so, why would he (she) treat me so unfairly?" They begin to ask: "What's in it for me?" and "Am I getting all I deserve in this relationship?" All would agree that when couples are at the point of breakup and divorce, they often become consumed with issues of profit (rewards minus costs) and fairness and equity.

Scholars agree that misery and unfairness are linked. They disagree, however, as to the nature of the causal relationship: Does perceived injustice cause dissatisfaction or is the causal order reversed? Clark (1986) takes the latter view. She argues that in communal relationships, couples do not "keep score;" they simply do not think in terms of reward and equity. Thus, if couples are concerned with such issues, it is a sure sign that their marriages are in trouble. Misery, then, is the *cause*, not the consequence of perceived injustice (Grote and Clark, 1998).

In a year-long longitudinal study, van Yperen and Buunk (1990) set out to answer this question. The authors interviewed couples who had been married for various lengths of time. They found that people in inequitable marriages became less satisfied over the course of a year. There was no evidence for the converse. It is possible, of course, that in failing marriages *both* processes are operating. In any case, it is clear that when marriages are faltering, people often become preoccupied with the pain and marital injustices they have endured, and this may well lead to relationship dissolution.

In sum, scientists have explored the impact that degree of reward and perceived fairness have on men and women's marital happiness and stability. It appears that although the concern with fairness may wax and wane during the course of a marriage, such concerns always remain there, just beneath the surface, guiding people's perceptions, happiness, and marital choices. Love is not blind.

[p. 205 ↓]

Ideas and Intellectual History of Equity Theory

Philosophical Underpinnings

Philosophers and ethicists have long been interested in the nature of social justice, fairness, and equity. The notion that others should be treated fairly is a ubiquitous one: more than 21 religions endorse some variant of the golden rule: "Do unto others ..." Philosophers Thomas Hobbes, John Locke, and Jean-Jacques Rousseau, spoke of the "social contract," Immanuel Kant of the "categorical imperative." John Rawls, a contemporary philosopher, speaks of "the veil of ignorance." All of these concepts ask "What is fair?" and propose that it is generally in a person's best interest to treat others as they deserve – at least most of the time.

Early Social Exchange Theorists

Social psychologists also attempted to understand the nature of social justice. The first modern-day scholars to propose models of social justice and social exchange (in the late 1950 and early 1960s) were sociologists George C. Homans (1958) and Peter Blau (1964) and social psychologists John Thibaut and Harold Kelley (1959). They viewed all social life as involving the exchange of goods, such as approval, esteem, and material goods. All people, they contended, are seeking maximum reward at minimum cost. As a consequence, there tends to be a balance in exchanges. Relying on the economic and behavioristic theories of the day, these scholars attempted to describe the factors that mediate the creation, maintenance, and breakdown of exchange relationships. Later, J. Stacy Adams (1965), an industrial-organizational psychologist, argued that management and labor attempt to maintain equity between the inputs they contribute to their work (such things as time, effort, and sacrifice) and the outcomes they receive from their professions (such things as security, salary, fringe benefits, etc.), compared

with the inputs and outputs of their supervisors, peers, and employees. Relational satisfaction was assumed to depend on how fair or unfair were distributions of rewards.

In the 1960s and 1970s, such social exchange theories were well received – so long as their advocates stuck to trying to predict behavior in academic, legal, or industrial settings. When we suggested that people might also be interested in reward and fairness in the arena of love, however, there was an explosion of indignation.

Early Equity Theorists and Critics

As we observed earlier, equity theory appeared in an era in which traditional views of gender roles, men's and women's liberation, and the rules of love and sex (including innovations such as marriage contracts) were being hotly debated. Thus, it is not surprising that the contention that couples care about "What's in it for me?" and "Am I being treated fairly" sparked a great debate. True, a prominent group of theorists – such as Peter Blau, in *Exchange and Power in Social Life*; Michael McCall, in *Courtship as Social Exchange*; Mirra Komarovsky, in *Blue-Collar Marriage*; Gerald R. Patterson, in *Families*; John Scanlon, in *Sexual Bargaining*; and Norman W. Storer, in *The Social System of Science* – agreed that in most intimate relationships, people do indeed care about pleasure and pain, fairness, and equity.

Yet, many people found the idea of speculating about the importance of equity in love relationships offensive. They harked back to Erich Fromm's writings. In *The Art of Loving*, Erich Fromm (1956: 3) declared that while flawed "human love relationships [may] follow the same pattern of exchange which governs the commodity and labor market, the truest form of love is unconditional love (love given without any thought of return"). He further observed that while men might be so crass as to act in selfish ways in love relationships, it was in women's nature to love unconditionally – giving without any [p. 206 ↓] thought of return. He notes, for example: "Fatherly love is conditional love. Its principle is 'I love you because you fulfill my expectations, because you do your duty, because you are like me ...' Motherly love is by its very nature unconditional" (1956: 35–36).

In the 1970s, a variety of social commentators agreed with Fromm's contention that people, especially women, are generally *not* concerned with reward or fairness in their love relationships – most notably Margaret Clark and Judson Mills (1979), Elizabeth Douvan (1974), and Bernard Murstein (Murstein et al., 1977).

Shortly after the publication of *Equity Theory and Research*, for example, Elaine was invited to give a speech at Yale. At Stanford, Minnesota, and Wisconsin, she was quite used to academic discussions of research. She was stunned when at Yale she received shouts and catcalls from the audience. One fellow, for example, complained that modern-day women were selfish and unfeminine in their concerns about marital fairness. Another shouted that his sainted mother had dedicated her life to him, with nary a thought of reward. "Pleasing the family," he contended, "should be woman's chief joy." Elaine was surprised that in that day and age, academics wouldn't be more aware of the complexities of human experience and of changing gender relationships. Certainly many wives and mothers are altruistic – but even they have moments when they wonder: "How did I get into this?" Certainly, anyone parenting a teenager thinks: "Wait until you have children of your own ..." Elaine had just talked to an elderly woman, who had nursed her sick husband for 30 years. After his death, she remarried. At first, she had been joyous, thinking: "Now is my time." When she discovered a few weeks later that her new husband was ill, and she was to be consigned to be a full-time nurse yet again, she couldn't help but cry out: "It isn't fair."

Clark (1986) argued that people participate in two kinds of relationships – communal relationships and exchange relationships. She observed:

In communal relationships, often exemplified by friendships and romantic relationships, people feel a special responsibility for one another's welfare. They give benefits in response to the other's needs or to please the other. In exchange relationships, often exemplified by acquaintances and business relationships, people feel no special responsibility for other's welfare. They give benefits with the expectation of receiving comparable benefits in return or in response to benefits previously received (1986: 414).

(Today, a number of theorists would agree with Clark's thesis that if couples are overly concerned with moment-to-moment equity, it is a "tip off" that there is something wrong with their relationships. See, for example, Aron et al., 1991; Sprecher, 1989; Van Lange et al., 1997.)

In a series of studies, Clark studied behavior in communal versus exchange settings. In these prototypic studies, a *communal orientation* was manipulated by introducing college men to an attractive *single* woman who acted as if she were interested in friendship. An *exchange orientation* was manipulated by introducing college men to an attractive *married* woman, who claimed to possess all the friends she desired. Men and women were assigned to work on some puzzles. During the encounter, the young woman asked the young man to assist her with her puzzles, and he complied. She then offered (or did not offer) to reciprocate. Men's feelings for the woman when she immediately offered (did not offer) to assist *him* depended on his orientation. In the communal condition, men preferred women who accepted help without immediately offering to pay them back. In the exchange setting, men preferred the (married) woman who accepted his aid, then offered to reciprocate in kind.

Clark (1986) and Williamson and Clark (1989) concluded that in dating, marital, and family relationships, communal norms prevail: men and women wish to please their partners, to care for and nurture them, and reject such crass considerations as "score-keeping" [p. 207 ↓] or a concern with *quid pro quo*. Relationships are complex, however, and a more cynical interpretation of Clark's results is possible. Normally it takes time for people to fall in love and commit themselves to an intimate relationship. In Clark's studies, men and women had just met. When "Prince Charming" assisted the "damsel in distress," there might have been two reasons why he preferred the attractive single women who did not insist on reciprocating in kind: (1) men might have possessed a communal orientation, as Clark believes; or (2) men may offer dinners, theatre tickets, and assistance to a beautiful woman, in hopes that she will willingly repay them with affection, gratitude, a date, or sex. The breathless, "How can I ever repay you?" is a TV cliché. In their heart of hearts, men in Clark's (1986) study may have been hoping to participate in an exchange – albeit a complex one.

On the face of it, the Clark perspective seemed diametrically opposed to our own. When one looks closer, however, the two often seem to merge. As they say, "The devil is in the details." Consider the following observations by Clark and her colleagues:

- Men and women prefer physically attractive mates, in part because the attractive are perceived to be more sensitive, kind, and capable of communal relationships than their peers (Clark, 1986).
- People who sacrifice on their partner's behalf, assume that their partners will be grateful, and become more loving and trusting than before, and thus more likely to "be there for *them* when the need arises" (Clark and Grote, 1998; Grote and Clark, 2004).
- Couples may prefer communal relationships, yet when desires and needs conflict, as they inevitably do, in the interests of fairness, men and women often decide to take turns in reaping benefits or suffering costs (Clark and Grote, 1998).
- People may differ in how communally oriented they are. A wife may assume her chivalrous husband is delighted to cater to her needs; her less communally oriented husband may resent what he considers to be her "exploitative" behavior (Mills et al., 2004).
- Some people are cunning and devious. A young medical student may ask his wife to put him through graduate school, only to divorce her upon graduation. In such cases, his communally oriented wife would naturally feel resentful at the betrayal (Williamson and Clark, 1989).
- When people suspect their mate is not communally oriented – that is, he does not care about their desires and needs – they will begin to mistrust the other, to "keep records," and worry about whether or not they are being fairly treated (Clark and Grote, 1998).

No matter how good a relationship, then, even in Clark's paradigm, it appears that now and then people ask themselves "Am I loved?" "Is my dating relationship or marriage rewarding?" "Is it fair and equitable?" The answers people come up with may well have a profound impact on their feelings about their relationships.

Development and Growth of Equity Theory over the Years

In many ways, the changing interests of equity theorists has paralleled changes in social psychology theorizing and research (in general) over the last 25 years (see Berscheid, 1992). Specifically, with the passage of time we became increasingly interested in the cultural, social, and biological forces that shape people's lives; a broadening and deepening of equity theory (to investigate more long-term relationships); and an increased sophistication in social psychological research methods.

Crafting Equity Theory

First, although from the start we yearned to integrate the insights of Darwinian theory, economic theory, and Hullian and Skinnerian reinforcement theories in crafting equity theory, in fact in those early days we were forced to focus more on nurture than nature. True, in the 1960s and 1970s, some pioneers like Hamilton (1964), Trivers (1971), and Smith (1974) assumed that altruism, as well as aggression, was wired into humankind. Theorists talked about the advantages of "group selection," "kin selection or inclusive [p. 208 ↓] fitness," and "reciprocal altruism" (a version of "If you scratch my back, I'll scratch yours"). Nonetheless, the influential theorist was Dawkins (1976), who contended in *The Selfish Gene* that day-to-day people are programmed for savage competition, ruthless exploitation, and deceit. Admittedly, altruistic acts occurred – but alas, such altruism was more apparent than real. Our challenge was, then, to craft a theory that accounted for people's desire for fairness and justice using primarily social constructionist and reinforcement models. Propositions I through IV focused on the social forces that prod people to care about social justice. The evidence for our contentions came, for the most part, from cultural psychology, social psychology, and Industrial-Organizational research.

In a subsequent section, we will discover that Darwinian theory has itself evolved. Today, evolutionary theorists have no trouble accounting for people's desire for fairness

and justice. Were we crafting equity theory today, we would be able to argue that people are predisposed both by nature and nurture to care very much indeed about social justice and to provide compelling evidence from a variety of disciplines for that contention.

Broadening and Deepening Equity Theory

In the early 1960s, I (EH) and my colleague Ellen Berscheid were primarily interested in theorizing about a totally neglected area – passionate love affairs. Our interest in this research topic was a natural one. First, we were young and all our friends were personally interested in this topic. Better yet, we were working in an as yet untrodden territory. Anything we discovered was bound to be interesting! Luckily, college students, often in the throes of passionate infatuations, were readily available as research participants.

The critics of equity theory – such as Margaret Clark and the late Judson Mills – were focused primarily on close, intimate, long-term relationships. Their contention that equity processes were likely to operate in a completely different way in deeply intimate settings was an exciting impetus to broaden our theory. It motivated us to try to deal with people's perceptions of fairness at every stage of the life cycle and to learn more about how equity processes played themselves out in these complex relationships.

The Appearance of Innovative Psychometric and Research Techniques

Since the 1970s, social psychologists have been hard at work developing a panoply of “user friendly” psychometric techniques for measuring psychological constructs. Technically, equity is defined by a complex formula (Traupmann et al., 1981; Walster, 1975). Respondents' perceptions as to the equitableness of their dating relationships or marriages is computed by entering their estimates of the inputs and outcomes of Persons A and B (/A

A

, I
B

, O
A

, and O
B

) into the equity formula:³

$$\frac{(O_A - I_A)}{(|I_A|)^{KA}} = \frac{(O_B - I_B)}{(|I_B|)^{KB}}$$

Respondents are classified as “over-benefited” if their relative gains exceed those of their partners. They are classified as “equitably treated” if their relative gains equal those of their partners, and as “under-benefited” if their relative gains fall short of those of their partners.

After conducting a great deal of research – and suffering through laborious data entries and complex calculations (which were often done by hand in that era), we abandoned our devilishly complex technique for assessing equity. Our original equity measure required us to ask questions (about inputs and outcomes) that respondents found it difficult to answer, and for the experimenter to record data, and perform calculations that were often difficult to perform – at least without errors! Then one day, it suddenly occurred to us that the best way to find out what we wished to know was to directly ask our respondents! Thus, in the Equity Global Measure (described earlier) we simply asked couples to think about their relationships and tell us how fair and equitable they seemed to be. Possible answers ranged from +3: “I am getting a much better deal than my partner” to - 3: “My partner is getting a much better deal than I am.”

[p. 209 ↓]

So obvious – and yet it had taken us a year or two to realize that this was the best way to find out what we wanted to know. (Sort of like the NASCAR builder who suddenly

realizes that the horse-and-buggy might not be the best automotive model.) This “time saving” was especially useful when we used the Multi-trait Measure of Equity. Twenty-five questions instead of 100!

Since the early days, social psychologists’ repertoire of research techniques has increased markedly. In the 1960s (as today), I (EH) was a specialist in social psychological experiments. Though I still value true experiments, now and then I and my colleagues turn to more complex methodologies – qualitative analyses, cross-sectional and long-term longitudinal studies, interactive paradigms, analyses of papal and church edicts, demographic data, historical and anthropological research, and the like.

The “hardware” of social psychologists has improved, too. Today, scholars can study people’s reactions to inequity by utilizing audiovisual recordings, social-psycho-physiological measures, fMRIs, physiological recordings, and a variety of unobtrusive measures. All of these new techniques give us a better understanding of the way concerns with Equity play themselves out in real life settings.

Current Research: A Multidisciplinary Approach

At the present time, some of the most interesting research into the nature of social justice emanates from scholars from four different intellectual traditions: (1) cultural scholars interested in societal definitions as to what is fair and equitable; (2) evolutionary theorists, who argue that a concern for justice arose early in humankind's long prehistory, and speculate about the ways in which this ancient “wiring” might affect contemporary visions of social justice; (3) neuroscientists, who are interested in charting the brain activity associated with perceptions of fairness or unfairness; and (4) primatologists, who speculate about the extent to which primates' sense of justice is similar or different from that of humankind.

Let us now consider a scattering of research from these four areas, just to give readers a sense of what is going on.

Is morality relative to culture? I would say that it is – and also that it isn't. Eric Knickerbocker

Equity: Cultural Considerations

Cultural theorists have long been interested in the impact of culture on perceptions of social justice. Anthropologists like Richard Shweder (1987) and Alan Fiske (2002), for example, surveyed moral concerns across the globe. All people, they argue, possess a sense of fairness; they assume people should reciprocate favors, reward benefactors, and punish cheaters.

Cultural theorists also contend that culture exerts a profound influence on how fairness is defined, how concerned men and women are that their intimate affairs be equitable, and how rewarding and equitable love relationships are likely to be (Amir and Sharon, 1987; Aumer-Ryan et al., 2007; Murphy-Berman and Berman, 2002).

Triandis and his colleagues (1990), for example, argued that in individualistic cultures (such as the US, Britain, Australia, Canada, and the countries of northern and western Europe) people generally focus on personal goals. In such societies, people are concerned with how rewarding (or punishing) their relationships are and how fairly (unfairly) they are treated. Collectivist cultures (such as China, many African and Latin American nations, Greece, southern Italy, and the Pacific Islands), on the other hand, insist that their members subordinate personal goals to those of the group: the family, the clan, or the tribe. It is tradition, duty, and deference to elders that matter. Rosenblatt and Cunningham (1976) claimed that equity is of less importance in collectivist societies: “[regardless of] who has the better life, a man or a woman, they [people of non-US cultures] might argue ... that the lives of men and women are different and not comparable” (cited in Buunk and VanYperen, 1989: 82).

[p. 210 ↓]

Do cultures differ in how much importance they attach to dating and marital fairness and equity? In a series of studies, Aumer-Ryan and her colleagues (2006) attempted to find out. They interviewed Japanese-American, West Indian, and multicultural internet users,

seeking answers to three questions. In different cultures, do men and women: (1) differ in the value they ascribe to equity in dating and marital relationships – some considering it to be crucial, others dismissing “fairness” as of trivial importance; (2) differ in whether they consider their own relationships to be equitable or inequitable; and (3) differ in how satisfied (or upset) when they discover their own relationships have turned out to be strikingly equitable/inequitable?

Aumer-Ryan and her colleagues (2006) found that, in all cultures, people considered reward and equity to be the gold standard of good relationships. Both Westerners and their non-Western counterparts insisted it was “important” or “very important” that a courtship relationship or marriage be equitable.

The authors did observe some fascinating cultural differences, however. People around the world may aspire to social justice, but few were lucky enough to achieve that goal. People in the various cultures differed markedly in how fair and equitable they considered their relationships to be. Men and women from the US claimed to be the most equitably treated. Men and women (especially women) from Jamaica, in the West Indies, felt the least equitably treated. Jamaican women often complained about men treating women as “second class citizens” and about men’s lack of commitment to relationships. In describing men’s attitudes, one woman quoted a classic Calypso song by Lord Kitchener (1963), and the repeated lyric: “You can always find another wife/ but you can never get another mother in your life.” Such attitudes, the women claimed, make it very difficult for them to find a relationship that is rewarding, fair, and fulfilling.

In all cultures, men and women reacted much the same way when they felt fairly or badly treated. All were most satisfied when receiving exactly what they felt they deserved from their relationships – no more (perhaps) but (just as in the West) certainly no less. (For additional research on this topic, see Murphy-Berman and Berman, 2002; Westerman, et al., 2007; Yamaguchi, 1994.)

Cultural psychologists not only give us information as to cultural differences in the way people define social justice. Cultural psychologists and historians also provide a window onto understanding the impact of social change on societal definitions of fairness. Some examples. Historians point out that globalization carries with it profound transformations in men and women’s roles and in gender equality (Hatfield and Rapson, 2005). What

impact does the movement from a traditional to a modern-day society have on the way men and women define fairness? How contented are men and women facing such changes? Do men tend to cling to the past while women rush into the future? How do men and women attempt to deal with the changes they perceive? Such research may provide new insights into the nature of social justice.

Is it not reasonable to anticipate that our understanding of the human mind would be aided greatly by knowing the purpose for which it was designed? (George C. Williams)

Equity: The Evolution of a Cultural Universal

In the past 25 years or so, social psychologists have begun to explore the evolutionary underpinnings of social justice. (See, for example, the classic work of Robert Trivers [1972] or Richard Dawkins [1976], on the probable evolution of reciprocal altruism and social exchange.) As Cosmides and Tooby observe:

It is likely that our ancestors have engaged in social exchange for at least several million years ... Social exchange behavior is both universal and highly elaborated across all human cultures – including hunter-gatherer cultures ... as would be expected if it were an ancient and central part of human life (1992: 164)

Currently, interesting work on social justice from evolutionary perspective is being [p. 211 ↓] conducted by scholars such as Rob Boyd (Boyd et al., 2003). They provide strong support for the notion in Proposition II that groups will reward those who treat others fairly and punish those who do not, even at considerable cost to themselves. Later in this chapter, we will discuss some of the evidence from neuroscientists and primatologists relevant to their observations.

Equity: FMRI Research

In recent years, neuroscientists have begun to investigate the cognitive factors (and brain processes) that are involved when men and women confront moral dilemmas. These concern such things as the nature of social justice and how a variety of competing moral claims – such as, “What's more important: the claims of friendship or the demands of fairness and equity in a social exchange?” – are resolved. Robertson and her colleagues (2007), for example, presented men and women with several real-life moral dilemmas. Using functional magnetic resonance imaging (fMRI) techniques, they studied people's brain activity as they pondered such dilemmas. The neuroscientists found that sensitivity to moral issues (in general) was associated with activation of the polar medial prefrontal cortex, dorsal posterior cingulated cortex, and posterior superior temporal sulcus (STS). They speculated that moral sensitivity is probably related to people's ability to retrieve autobiographical memories and to take a social perspective. They also assessed whether sensitivity to social concerns as distinguished from impartial justice involved different kinds of neural processing. They found that sensitivity to issues of justice (and social exchange) was associated with greater activation of the left intraparietal sulcus, whereas sensitivity to care issues was associated with greater activation of the ventral posterior cingulate cortex, ventromedial and dorsolateral prefrontal cortex, and the thalamus. These results suggest that different parts of the brain may operate when people ponder their duty to loved ones versus their obligation to be fair and just to all. For additional neurobiological speculations as to the neural circuits involved in the perception of and reaction to social inequality, see Borg et al. (2006), Raine and Yang (2006), Reis et al. (2007), Watson and Platt (2006), and Witvliet et al. (2008).

Neuroscience is still in its infancy, of course. Many social scientists have sharply criticized the widespread use of fMRI techniques to study the nature of social justice, claiming that currently the fMRI studies track only superficial changes and lack reliability and validity (Cacioppo, et al., 2003; Movshon, 2006; Panksepp, 2007; Wade, cited in Wargo, 2005). Nonetheless, this groundbreaking research has the potential (as it grows ever more sophisticated) to answer age-old questions as to the nature of culture,

perceptions of social justice, and the ways in which people react when faced with equitable or inequitable treatment.

Equity: The Concern for Justice in other Species

Today, paleoanthropological evidence supports the view that notions of social justice and equity are extremely ancient. Ravens, for example, have been observed to attack those who violate social norms. Dogs get jealous if their playmates get treats and they do not. Wolves who don't "play fair" are often ostracized – a penalty that may well lead to the wolf's death (Bekoff, 2004; Brosnan, 2006).

Primateologists have amassed considerable evidence that primates and other animals do care about fairness. In a study with brown capuchin (*Cebus apella*) monkeys, Brosnan and de Waal (2003) discovered that female monkeys who were denied the rewards they deserved became furious. They refused to play the game (i.e., refused to exchange tokens for a cucumber) and disdained to eat their prize – holding out for the grapes they deserved. If severely provoked (the other monkey did nothing and still got the highly prized grapes instead of the cucumber) capuchins grew so angry that they began to scream, beat their breasts, and hurl food at the experimenter. Interestingly, in a later study, the authors found [p. 212 ↓] that chimpanzees (*Pan troglodytes*) were most upset by injustice in casual relationships. In chimps' close, intimate relationships, injustice caused barely a ripple (Brosnan et al., 2005). There is some evidence that in close groups, chimps will voluntarily set things right. We see, then, that different species, in different settings, may respond differently to injustice. (Some critics have argued that in these experiments there is a confound between "impaired self-interest" and "injustice," since injustice was manipulated by denying the chimps reward. Only subsequent research can determine whether or not these primates can truly be said to be seeking justice.)

In the late 1990s, Ronald Noë and Peter Hammerstein (1994) proposed the notion of "biological markets" – predicting that primates would respond to much the same market forces as men and women do in selecting their mates. Recently, Ronald Gumert (2008)

observed 50 longtailed macaques in a reserve in Central Kalimantan, Indonesia. He found that in essence male macaque monkeys pay for sex by grooming the female. "It suggests that sex is a commodity," he observed. And as with other commodities, the value of sex is affected by considerations of supply and demand. Gumert observes:

A male would spend more time grooming a female if there were fewer females in the vicinity ... And when the female supply is higher, the male spends less time on grooming ... The mating actually becomes cheaper depending on the market (2008: 1)

Potentially, this fascinating animal research may provide some insights into three questions that have intrigued equity researchers:

- When, in primates' long prehistory, did animals begin to feel "guilty" about receiving "too much," as well as feeling outraged when they are "ripped off"? (Brosnan et al., 2005; Brosnan, 2006).
- Are animals more (or less) concerned about fairness in despotic, hierarchical societies than in those that are relatively egalitarian? (Brosnan, 2006).
- Are primates and other animals more (or less) concerned about inequities in close kin relationships than in more distant encounters? (Brosnan et al, 2005).

Equity Theory's Applicability to Relationship Issues

Essentially, the equity argument goes as follows: People may be motivated by self-interest, but they soon learn that the best way to survive and prosper is by following social rules as to what is fair or unfair. Thus, all in all, men and women will feel most comfortable when they are getting roughly what they deserve from their relationships. This fact has several practical implications:

Mate Selection

When pop-psychology authors give advice, they often assume that all their readers are entitled to all of life's riches. Romantics are eager to take such "advice." We once asked one of our clients, who complained that all the good men were taken, what she was looking for in a mate. She quickly scribbled a list of qualities she considered indispensable. Her list comprised more than 200 items! Many of them were contradictory. ("I want an ambitious and successful moneymaker" and "I want someone who will do at least 50 percent of the housecleaning and childcare.") This list was presented without a trace of irony. Many people, sad to say, think very much along these lines and are convinced that through the magic of positive thinking, "affirmations," <http://eHarmony.com>, and such, they can "have it all" (Rapson, 2008). And, you may ask, why not? To equity theorists, such expectations are wildly impractical. Of course everyone longs for perfection. Unfortunately, the supply of perfection is somewhat limited.

In the musical *Showboat*, a hard working showman, Frank, pleads with a young girl, Ellie, to marry him. Her saucy reply:

After I have looked around the world for a mate, Then, perhaps, I might fall back on you! When I am convinced that there is no better fate, Then I might decide that you will do.

A harsh way of putting it, but indeed there is more than a grain of truth in her soliloquy. [p. 213 ↓] People do often consider their "market value" when deciding what they will settle for in a mate. In real life, imperfect humans with run-of-the-mill flaws (like all of us) had better resign themselves to the fact that they will have to settle for other humans no better and no worse than themselves.

Saints and Sinners

How do couples, caught up in unbalanced relationships, generally handle their feelings of distress? Equity theorists observed that men and women can reduce distress via three techniques:

- 1. *Restoration of actual equity.* One way individuals can restore equity to an unjust relationship is by voluntarily setting things right, or by urging their partners to do so. Couples do often make considerable effort to balance things out. The husband who has been irritable because of stress at work may try to make amends by taking the family on a holiday when the pressure lets up.
- 2. *Restoration of psychological equity.* Couples in inequitable relationships can reduce distress in a second way. They can distort reality and convince themselves (and perhaps others as well) that things are perfectly fair just as they are. A variety of studies have documented the imaginative techniques that people use to justify injustice. Some studies find, for example, that harm-doers rationalize the harm they inflict on others by denying they are responsible for the victim's suffering ("I was just following orders"), by insisting that the victim deserved to suffer, or by minimizing the extent to which the victim suffered from their actions. There is even some sparse experimental evidence that, under the right circumstances, victims will justify their own exploitation.
- 3. *Leave the relationship.* Finally, if couples are unable to restore equity to their intimate relationships, there is a third way they can try to set things right. They can leave the relationship. This does not always mean divorce. A person will sometimes "opt-out" by abandoning their partners emotionally. New mothers, less attracted to their husbands than to their newborns, may insist that their infants sleep between them. This is a most effective strategy for keeping the couple apart. Or couples may spend all their leisure time "drinking with the boys" or "shopping with the girls," ensuring that they will rarely spend time alone together as a twosome. Both partners may risk their hearts in extramarital affairs. Or, finally, they may simply leave altogether.

The vast literature on how people deal with inequity has practical implications for close relationships. We know, for example, that people come to love those they treat with kindness and to despise those they abuse. Relationships should go best when they are balanced, when both people love one another, sacrifice for one another, and are loved in return.

Yet, even in the best of relationships, people have to be wary if they spot things going awry. If you think about the close relationships of some of your friends, you can probably come up with some examples of people who are emotionally stingy and hence give too little in relationships or who are always willing to give too much. How have these unbalanced affairs have worked out? Here are a couple of examples from our experiences as psychotherapists:

One of our clients was appallingly narcissistic. He was good-looking and had a sort of raffish charm, but he wasn't willing to make *any* compromises. "You compromise once," he said, "and you set a precedent; there's no end to it." In singles bars, women swarmed around him. However, once they started spending time with him, they soon became irritated. At first they could convince themselves that it was "just this once" that they would be stuck in the kitchen preparing a spontaneous dinner for ten of his pals while they watched the Super Bowl. Only this once would he ask her to research and type his reports while he took a nap. But as the days turned into weeks and the "just-this-onces" became a mantra, the rationalizations turned to seething rage. The women felt ripped off. Eventually they left the kitchen and the computer keyboard, and walked away from the relationship. If men and women know deep down that they are taking advantage of their partners, they should be warned that they may be playing a dangerous game. Sometimes people win all the battles only to lose the war. Their partners give in and give in, until finally they have had enough: they get fed up and leave.

[p. 214 ↓]

Another of our clients was very paternal; he was always attracted to "wounded birds" – beautiful young women who were so troubled and so uneducated that they couldn't make it on their own. He tried to anticipate all their needs and showered them with expensive presents. Any time that trouble threatened, he tried harder and gave yet

more. Inevitably, his relationships fell apart. His young girlfriends were grateful; they felt they *should* love him (and were ashamed that they couldn't). But they just didn't. "Where was his self-respect? Why was he so desperate?" "What was wrong with him?" They felt smothered. They couldn't bear to touch him. They had to flee.

Men and women have to be able to set limits. If loving people become aware that their mates are taking them for granted or treating them like doormats, they must have the strength to complain, draw the line, or give up the relationship. Otherwise, the relationship becomes a dangerously inequitable one.

Equity Theory's Applicability to Social Issues

In this chapter, we come full circle. When discussing the origins of equity theory, we spoke of the social ferment of the 1960s and 1970s. Today, 40 years later, in the West, these spirited debates about the nature of social justice have borne fruit. As Rapson (2008) observed, scholars generally agree that the two most significant social transformations of the twentieth and twenty-first centuries are: (1) the globalization of science and technology; and (2) the women's movement.

In the West, we have not yet achieved gender equality, though the approach toward that destination has been rapid in recent decades, particularly in parts of northern Europe. Male supremacy, however, continues to be the rule worldwide. When the United Nations sponsored a trio of human rights conferences in Vienna, Geneva, and Beijing, members of the world community agreed that abuses of women's rights – which include female infanticide, genital mutilation, the sale of brides, dowry murders, *suttee* or widow burning (in India, widows are still sometimes required to immolate themselves on their husbands' funeral pyres), and discriminatory laws against women's civic, social, and legal equality – are, in fact, abuses of human rights generally. The members also agreed that "culture" and "tradition" can no longer be cited to justify repression of half the world's population, especially since "tradition" has been defined by the powerful men in those societies. Gender equality around the world, including the developed world, is a long way off.

Yet, the winds of modernism are sweeping even into the most well-guarded sanctuaries of the male privilege. A revolution in gender and love relations has begun in China, Japan, Mexico, Latin America, North Africa, Russia, and even in Turkey, Saudi Arabia, and some other Arab nations (see Hatfield and Rapson, 2005, for the research in support of these conclusions). In different societies, these historic changes are moving at different speeds, of course. The women's movement has had a greater impact on men and women's lives in Europe than in the Middle East, in urban than in rural settings, in secular societies than in theistic cultures, and in wealthier than in poorer nations. But the times they are a'-changin' and the changes we have described are of monumental significance.

These epochal historical movements are likely to profoundly transform men and women's perceptions as to what is fair and equitable in love relationships. Historically, women have had little power. This meant that they possessed little freedom to shape their own lives, were forced to resign themselves to minimal expectations, and to be dependent on husbands and sons for most of life's basic needs. But as women gain more social power, they are likely to possess more bargaining power, higher expectations, and be more demanding in the arena of life and love. Given these social changes – as surely as the earth orbits the sun – women's demands for fair and equitable treatment are sure to grow and, consequently, so will men's. We believe that the continued global march towards gender equality will enlarge demands for [p. 215 ↓] equity in love relationships and profoundly alter (mostly for the better, we believe) love relationships around the globe.

Conclusion

We have traced equity theory from its beginnings to the present day. We reviewed compelling evidence that at all stages in love relationships – from their tentative beginnings, through their flowering, and perhaps to their bitter ending – men and women are concerned with both reward and fairness. We reviewed new multidisciplinary research – cultural, evolutionary, primatological, and neuroscience investigations – which add new depth and richness to our understanding of human nature. We closed by pointing out that the massive social changes that are occurring in our times, suggesting

that men and women may well be developing more complex and (especially for women) more fulfilling notions as to what constitutes fair and equitable treatment in love and life.

Notes

1 As you can see, in developing this theory we were hoping to combine the insights of Darwinian theory, economic theory, and reinforcement theory.

2 And perplexity or disgust when contemplating their partner's weakness.

3 The equity formulas used by previous researchers, from Aristotle to Stacy Adams, only yield meaningful results if A and B 's inputs and outcomes are entirely positive or entirely negative. In mixed cases the formulas yield extremely peculiar results. Thus, we proposed an equity computational scheme designed to transcend these limitations. See Walster (1995) for a discussion of the problems and the mathematical solutions. The superscript K simply "scales" equity problems (by multiplying all inputs and outcomes by a positive constant) such that the minimum of $|/$

A

$|$ and $|/$

B

$|$ is greater than or equal to 1.

References

Adams, J.S. Inequity in social exchange . Advances in Experimental Social Psychology , 62, 335–343. (1965)

Amir, Y. and Sharon, I. Are social psychological laws cross-culturally valid? Journal of Cross-Cultural Psychology , 18, 383–470. (1987) <http://dx.doi.org/10.1177/0022002187018004002>

Anselm of Canterbury (1998) *Opera omnia* . In B. Davies, ed. and G. Evans (eds), *Anselem of Canterbury: The Major Works* . New York: Oxford University Press. (Original work 1070–1109 AD).

Aron, A., Aron, E.N., Tudor, M. and Nelson, G. Close relationships as including other in the self . *Journal of Personality and Social Psychology* , 60, 241–253. (1991) <http://dx.doi.org/10.1037/0022-3514.60.2.241>

Aumer-Ryan, K., Hatfield, E. and Frey, R. (2007) Equity in romantic relationships: An analysis across self-construal and culture . University of Texas, Austin, Texas.

Baumeister, R.F. and Vohs, K.D. Sexual economics: Sex as female resource for social exchange in heterosexual interactions . *Personality and Social Psychology Review* , 8, 339–363. (2004) http://dx.doi.org/10.1207/s15327957pspr0804_2

Bekoff, M. (2004) Wild justice, cooperation, and fair play: Minding manners, being nice, and feeling good . In R. Sussman, ed. and A. Chapman (eds), *The Origins and Nature of Sociality* , pp. 53–79. Chicago: Aldine.

Bernard, J. (1972) *The Future of Marriage* . New York: World.

Berscheid, E. A glance back at a quarter century of social psychology . *Journal of Personality and Social Psychology* , 63, 525–533. (1992) <http://dx.doi.org/10.1037/0022-3514.63.4.525>

Blau, P. (1964) *Exchange and Power in Social Life* . New York: Wiley.

Blau, P. (1986) *Exchange and Power in Social Life* . New York: Transition Press.

Borg, J.S., Hynes, C., Van Horn, J., Grafton, S. and Sinnott-Armstrong, W. Consequences, action, and intention as factors in moral judgments: An fMRI investigation . *Journal of Cognitive Neuroscience* , 18, 803–817. (2006) <http://dx.doi.org/10.1162/jocn.2006.18.5.803>

Boyd, R., Gintis, H., Bowles, S. and Richerson, P.J. The evolution of altruistic punishment . PNAS , 100, 3531–3535. (2003) <http://dx.doi.org/10.1073/pnas.0630443100>

Brosnan, S.F. At a crossroads of disciplines . Journal of Social Justice , 19, 218–227. (2006) <http://dx.doi.org/10.1007/s11211-006-0003-y>

Brosnan, S.F., Schiff, H.C. and de Waal, F.B.M. Tolerance for inequity may increase with social closeness in chimpanzees . Proceedings of the Royal Society of London Series B 1560, 253–258. (2005)

Brosnan, S.F. and de Waal, F.B.M. Monkeys reject unequal pay . Nature , 425, 297–299. (2003) <http://dx.doi.org/10.1038/nature01963>

Buunk, B.P. and Van Ypern, N.W. Social comparison, equality, and relationship satisfaction: Gender differences over a ten-year period . Social Justice Research , 3, 157–180. (1989) <http://dx.doi.org/10.1007/BF01048064>

Byers, E.S. and Wang, A. (2004) Understanding sexuality in close relationships from the social exchange perspective . In J.H. Harvey, ed. , A. Wenzel, ed. , and S. Sprecher (eds), Handbook of Sexuality in Close Relationships , pp. 203–234. Mahwah, NJ: Lawrence Erlbaum.

Cacioppo, J.T., Berntson, G.G., Lorig, T.S., Norris, C.J. and Nusbaum, H. Just because you're imaging the brain doesn't mean you can stop using your head: A primer and set of first principles . Journal of Personality and Social Psychology , 85, 650–661. (2003) <http://dx.doi.org/10.1037/0022-3514.85.4.650>

Clark, M.S. Evidence for the effectiveness of manipulations of communal and exchange relationships . Personality and Social Psychology Bulletin , 12, 414–425. (1986) <http://dx.doi.org/10.1177/0146167286124004>

Clark, M.S. and Grote, N.K. Why aren't indices of relationship costs always negatively related to indices of relationship quality? Personality and Social Psychology Review , 2, 2–17. (1998) http://dx.doi.org/10.1207/s15327957pspr0201_1

Clark, M.S. and Mills, J. Interpersonal attraction in exchange and communal relationships . *Journal of Personality and Social Psychology* , 37, 12–24. (1979) <http://dx.doi.org/10.1037/0022-3514.37.1.12>

Cosmides, L. and Tooby, J. (1992) Cognitive adaptations for social exchange . In J.H. Barkow, ed. , L. Cosmides, ed. , and J. Tooby (eds), *The Adapted Mind* , pp. 161–228. New York: Oxford University Press.

Dawkins, R. (1976) *The Selfish Gene* . Oxford: Oxford University Press.

Douvan, E. (1974) Interpersonal relationships — some questions and observations . Paper presented at the Rausch Conference, Durham, NC, Duke University.

Fiske, A.P. Moral emotions provide the self-control needed to sustain social relationships . *Self and Identity* , 1, 169–175. (2002) <http://dx.doi.org/10.1080/152988602317319357>

Fromm, E. (1956) *The Art of Loving* . New York: Harper & Row.

Fierstone, S. (1979) *The Dialectic of Sex* . New York: Women's Press.

Friedan, B. (1963) *The Feminine Mystique* . New York: W.W. Norton.

Goffman, E. On cooling the mark out: Some aspects of adaptation to failure . *Psychiatry* , 15, 451–463. (1952)

Grote, N.K. and Clark, M.S. Distributive justice norms and family work: What is perceived as ideal, what is applied, and what predicts perceived fairness . *Social Justice Research* , 11, 243–269. (2004) <http://dx.doi.org/10.1023/A:1023234732556>

Gumert, M. (2008) Macaque monkeys 'pay' for sex . *New Scientist* , January, 1–2.

Hamilton, W.D. The genetical evolution of social behaviour I and II . *Journal of Theoretical Biology* , 7, 1–32. (1964) <http://dx.doi.org/10.1016/0022-5193%2864%2990038-4>

Hatfield, E. and Rapson, R.L. (1993) Love, Sex, and Intimacy: Their Psychology, Biology, and History . New York: Harper/Collins.

Hatfield, E. and Rapson, R.L. (2005) Love and Sex: Cross-Cultural Perspectives . Lanham, MD: University Press of America.

Hatfield, E., Rapson, R.L. and Aumer-Ryan, K. (2007) Equity Theory: Social justice in love relationships. Recent developments . Social Justice Research . New York: Springer.

Hatfield, E., Rapson, R.L. and Aumer-Ryan, K. (2008) Social justice in love relationships: Recent developments . Social Justice Research .

Hatfield, E., Walster, G.W. and Berscheid, E. (1978) Equity: Theory and Research . Boston: Allyn & Bacon.

Homans, G.C. Social behavior as exchange . American Journal of Sociology , 63, 97–606. (1958)

Kitchener, L. (1963) Mother and wife [Recorded by The Invaders Steel Band] . On Air Mail Music: Steel Bands Caraibes [CD]. Track 12 . Boulogne, France: Playasound.

Komarovsky, M. (1964) Blue-Collar Marriage . New York: Random House.

Lawrance, K-A. and Byers, E.S. Sexual satisfaction in long-term heterosexual relationships: The interpersonal exchange model of sexual satisfaction . Personal Relationships , 2, 267–285. (1995) <http://dx.doi.org/10.1111/j.1475-6811.1995.tb00092.x>

Markman, H.J. Prediction of marital distress: A 5-year follow up . Journal of Consulting and Clinical Psychology , 49, 460–762. (1981) <http://dx.doi.org/10.1037/0022-006X.49.5.760>

Marsella, A.J. Toward a global psychology: Meeting the needs of a changing world . American Psychologist , 53, 1282–1291. (1998) <http://dx.doi.org/10.1037/0003-066X.53.12.1282>

Martin, M.W. Satisfaction with intimate exchange: Gender-role differences and the impact of equity, equality, and rewards . *Sex Roles* , 13, 597–605. (1985) <http://dx.doi.org/10.1007/BF00287297>

Mills, J., Clark, M.S., Ford, T.E. and Johnson, M. Measurement of communal strength . *Personal Relationships* , 11, 213–230. (2004) <http://dx.doi.org/10.1111/j.1475-6811.2004.00079.x>

Movshon, J.A. (2006) Searching for the person in the brain . *The New York Times. Week in Review* , February 5, 1–4.

Murphy-Berman, V. and Berman, J. Cross-cultural differences in perceptions of distributive justice . *Journal of Cross-Cultural Psychology* , 33, 157–170. (2002) <http://dx.doi.org/10.1177/0022022102033002003>

Murstein, B.I., Cerreto, M. and MacDonald, M.G. A theory and investigation of the effect of exchange-orientation on marriage and friendship . *Journal of Marriage and the Family* , 39, 543–548. (1977) <http://dx.doi.org/10.2307/350908>

Noë, R. and Hammerstein, P. Biological markets: supply and demand determine the effect of partner choice in cooperation, mutualism, and mating . *Behavioral Ecology and Sociobiology* , 35, 1–11. (1994)

Panksepp, J. Neurologizing the psychology of affects: How appraisal-based constructivism and basic emotion theory can coexist . *Perspectives on Psychological Science* , 2, 281–312. (2007) <http://dx.doi.org/10.1111/j.1745-6916.2007.00045.x>

Patterson, G. R. (1971) Families: Applications of Social Learning to Family Life . Champaign, IL: Research Press.

Perlman, D. and Duck, S. (1986) Intimate Relationships: Development, Dynamics, and Deterioration . New York: Sage.

Pillemer, J., Hatfield, E. and Sprecher, S. The importance of fairness and equity for the marital satisfaction of older women . *Journal of Women and Aging* , 20, 215–230. (2008) <http://dx.doi.org/10.1080/08952840801984493>

Raine, A. and Yang, Y. Neural foundations to moral reasoning and antisocial behavior . *Social Cognitive and Affective Neuroscience* , 1, 203–213. (2006) <http://dx.doi.org/10.1093/scan/nsl033>

Rapson, R.L. (2008) *Magical Thinking and the Decline of America* . Philadelphia: Xlibris.

Reis, D.L., Brackett, M.A., Shamosh, N.A., Kiehl, K.A., Salovey, P. and Gray, J.R. Emotional intelligence predicts individual differences in social exchange reasoning *NeuroImage* , 35, 1385–1391. (2007) <http://dx.doi.org/10.1016/j.neuroimage.2006.12.045>

Robertson, D., Snarey, J., Ousley, O., Harenski, K., Bowman, F.D., Gilkey, R. and Kilts, C. (2007) The neural processing of moral sensitivity to issues of justice and care . *Neuropsychologia* , 755–766. <http://dx.doi.org/10.1016/j.neuropsychologia.2006.08.014>

Rosenblatt, P.C. and Cunningham, M.R. (1976) Sex differences in cross-cultural perspective . In B. Lloyd, ed. and J. Archer (eds), *Exploring Sex Differences* , pp. 71–94. London: Academic Press.

Scanzoni, J. (1972) *Sexual Bargaining: Power Politics in the American Marriage* . Englewood Cliffs, NJ: Prentice-Hall.

Schreurs, K.M.G. and Buunk, B.P. Closeness, autonomy, equity, and relationship satisfaction in lesbian couples . *Psychology of Women Quarterly* , 20, 577–592. (1996) <http://dx.doi.org/10.1111/j.1471-6402.1996.tb00323.x>

Shweder, R.A. (1987) Culture and moral development . In J. Kagen, ed. and S. Lamb (eds), *The Emergence of Morality in Young Children* , pp. 1–88. Chicago: University of Chicago Press.

Smith, M. (1974) *Models in Ecology* . Cambridge: Cambridge University Press.

Sprecher, S. The effect of exchange orientation on close relationships . *Social Psychology Quarterly* , 61, 220–231. (1989) <http://dx.doi.org/10.2307/2787109>

Storer, N.W. (1966) *The Social System of Science* . New York: Holt, Rinehart, and Winston.

Thibaut, J.W. and Kelley, H.H. (1959) *The Social Psychology of Groups* . New York: John Wiley.

Tooby, J. and Cosmides, L. (1992) The evolutionary and psychological foundations of the social sciences . In J.H. Barkow, ed. , L. Cosmides, ed. and J. Tooby (eds). *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* , pp. 19–136. New York: Oxford University Press.

Traupmann, J., Peterson, R., Utne, M. and Hatfield, E. Measuring equity in intimate relations . *Applied Psychological Measurement* , 5, 467–480. (1981) <http://dx.doi.org/10.1177/014662168100500405>

Triandis, H.C., McCusker, C. and Hui, C.H. Multimethod probes of individualism and collectivism . *Journal of Personality and Social Psychology* , 59, 1006–1020. (1990) <http://dx.doi.org/10.1037/0022-3514.59.5.1006>

Trivers, R.L. The evolution of reciprocal altruism . *Quarterly Review of Biology* , 46, 35–57. (1971) <http://dx.doi.org/10.1086/406755>

Trivers, R.L. The evolution of reciprocal altruism . *Quarterly Review of Biology* , 46, 35–37. (1972)

Van Lange, P.A.M., Rusbult, C.E., Drigotas, S.M., Ariaga, X.B., Witcher, B.S. and Cox, C.L. Willingness to sacrifice in close relationships . *Journal of Personality and Social Psychology* , 72, 1373–1395. (1997) <http://dx.doi.org/10.1037/0022-3514.72.6.1373>

van Yperen, N.W. and Buunk, B.P. A longitudinal study of equity and satisfaction in intimate relationships . *European Journal of Social Psychology* , 20, 287–309. (1990)

Walster, G.W. Equity formula: A correction . *Representative Research in Social Psychology* , 6, 65–67. (1975)

Wargo, E. With the brain, is seeing believing? *American Psychological Society* , 18, 33. (2005)

Watson, K. and Platt, M.L. Fairness and the neurobiology of social cognition: Commentary on 'Nonhuman species' reactions to inequity and their implications for fairness' by Sarah Brosnan . *Social Justice Research* , 19, 186–193. (2006) <http://dx.doi.org/10.1007/s11211-006-0008-6>

Westerman, C.Y.K., Park, H.S. and Lee, H.E. A test of equity theory in multidimensional friendships: A comparison of the United States and Korea . *Journal of Communication* , 57, 576–598. (2007) <http://dx.doi.org/10.1111/j.1460-2466.2007.00358.x>

Williams, G.C. (1996) *Adaptation and Natural Selection* . Princeton: Princeton University Press.

Williamson, G.M. and Clark, M.S. The communal/exchange distinction and some implications for understanding justice in families . *Social Justice Research* , 3, 77–103. (1989) <http://dx.doi.org/10.1007/BF01048060>

Witvliet, C.V.O., Worthington, E.L., Root, L.M., Sato, A.F., Ludwig, T.E. and Exline, J.J. Retributive justice, restorative justice, and forgiveness; an experimental psychophysiology analysis . *Journal of Experimental Social Psychology* , 44, 10–25. (2008) <http://dx.doi.org/10.1016/j.jesp.2007.01.009>

Yamaguchi, S. (1994) Collectivism among the Japanese: A perspective from the self . In U. Kim, ed. , H.C. Triandis, ed. and G. Yoon (eds), *Individualism and Collectivism: Theoretical and Methodological Issues* , pp. 175–188. Thousand Oaks: Sage.

Young, D. and Hatfield, E. (2009) Measuring equity in close sexual relationships . In T.D. Fisher, ed. , C.M. Davis, ed. , W.L. Yaber, ed. and S.L. Davis (eds), *Handbook of Sexuality-related Measures: A Compendium* , 3rd Edition. Thousand Oaks, CA: Taylor & Francis.

<http://dx.doi.org/10.4135/9781446249222.n36>

Handbook of Theories of Social Psychology

The Investment Model of Commitment Processes

Contributors: Caryl E. Rusbult & Christopher R. Agnew & Ximena B. Arriaga

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "The Investment Model of Commitment Processes"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n37>

Print pages: 218-232

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n37>

[p. 218 ↓]

Chapter 37: The Investment Model of Commitment Processes

Caryl E.Rusbult¹ Christopher R.Agnew Ximena B.Arriaga

Abstract

The investment model of commitment processes is rooted in interdependence theory and emerged from the broader scientific zeitgeist of the 1960s and 1970s that sought to understand seemingly irrational persistence in social behavior. The investment model was developed originally to move social psychology beyond focusing only on positive affect in predicting persistence in a close interpersonal relationship. As originally tested, the investment model holds that commitment to a target is influenced by three independent factors: satisfaction level, quality of alternatives, and investment size. Commitment, in turn, is posited to mediate the effects of these three bases of dependence on behavior, including persistence. Commitment is presumed to bring about persistence by influencing a host of relationship maintenance phenomena. The investment model has proven to be remarkably generalizable across a range of commitment targets, including commitment toward both interpersonal (e.g., abusive relationships, friendships) and noninterpersonal (e.g., job, sports participation, support for public policies) targets. Empirical support for the investment model is presented as well as a review of recent applications of the model and a proposed extension of it.

Introduction

The investment model (Rusbult, 1980, 1983) provides a useful framework for predicting the state of being committed to someone or something, and for understanding the

underlying causes of commitment. It was developed to move beyond focusing only on positive affect in predicting persistence in an interpersonal relationship. A major premise of the investment model is that relationships persist not only because of the positive qualities that attract partners to one another (their satisfaction), but also because of the ties that bind partners to each other (their investments) and the absence of a better option beyond the relationship with the current partner (lack of alternatives); all of these factors matter in understanding commitment. Beyond explaining the antecedents of commitment, the investment model has generated a large body of research to account for what differentiates lasting relationships from those that end and on specific cognitive and behavioral [p. 219 ↓] maintenance mechanisms that are fueled by commitment. The model also has been applied to predicting commitment to all sorts of other targets, revealing its generalizability beyond close relationships.

Origins of the Investment Model

In the summer of 1976, I (CER) took a crosscountry road trip from Chapel Hill to Los Angeles. I had just completed my first year of graduate school at the University of North Carolina at Chapel Hill (UNC), and was keen to visit friends and family in LA. On the return leg of the trip my traveling companion brought up an interesting topic: "Tell me why people stick with their partners." I spent the better part of Arizona and New Mexico describing work that seemed relevant – work regarding attitudinal similarity, physical appearance, the gain–loss phenomenon, pratfall effects, and the like. I gave a good review of the relationships literature as it existed in the mid 1970s. As we crossed the Texas border, however, my traveling companion somewhat sheepishly asked: "Okay, but can you tell me why people stick with their partners?" My companion was correct in his implied assessment of this literature. Although work regarding interpersonal attraction answers some interesting questions – for example, what makes us feel attracted to a partner, and what makes us feel satisfied with a relationship – it does not explain why people sometimes persist in relationships. The issue in this literature was positive affect, not perseverance.

As it turns out, several months earlier I had participated in a seminar on interdependence processes led by John Thibaut at UNC. Interdependence theory, which John developed with Hal Kelley (Thibaut and Kelley, 1959; Kelley and Thibaut,

1978), argues that dependence is a central structural property of relationships and particularly relevant to understanding persistence. We describe dependence in more detail below, but the point here is that there was a compelling theory suggesting that dependence, not satisfaction, drives people to seek further interaction with each other. In the context of ongoing romantic involvements, this meant that relationships persist not only as a function of the positive or negative qualities that derive from a particular partner, but also because being with the partner on the whole is more desirable than not being with the partner.

As described in Rusbult et al. (2006), the investment model was also shaped by the broader scientific zeitgeist in the 1960s and 1970s that sought to explain unjustified persistence. During this period, social scientists from diverse fields sought to understand “irrational persistence” in nonromantic domains. Social science research repeatedly documented commitment-relevant phenomena, such as dedicating more time or effort than desired to a particular activity; increasing commitment to a losing enterprise (i.e., the irrational escalation of commitment; Staw, 1976); being trapped in an escalating conflict (such as in the dollar auction, a bidding game in which a dollar goes to the highest bidder but the second-highest bidder must also pay the highest amount that he or she bid; Shubik, 1971); and the manner in which investments, side bets, and sunk costs may induce perseverance in a line of action (Becker, 1960; Blau, 1967; Brockner et al., 1979; Kiesler, 1971; Schelling, 1956; Teger, 1980; Tropper, 1972; see Rusbult, 1980). Increased scholarly interest in irrational persistence during that time was on par with the broader sociopolitical events, such as the Cold War arms race and US involvement in Vietnam. This is not to say that the investment model, or other models cited here, was directly inspired by such events, but rather that during this era, a fascination with unjustified persistence was “in the air” from a scientific point of view (Rusbult et al., 2006).

[p. 220 ↓]

Initial Tests of the Investment Model

With important personal and scholarly influences as a backdrop, Rusbult completed her dissertation in 1978, in which she developed the investment model. Initial tests of

the investment model were published in the early 1980s (Rusbult, 1980, 1983). These early papers included: (1) an experiment in which participants read one vignette of a hypothetical couple varying, in a between-subject design, the costs, alternatives, and investment size to assess the effect of these variables on satisfaction and commitment (Rusbult, 1980, Study 1); (2) a cross-sectional study in which participants completed a survey with respect to their own relationship, assessing the association of costs, rewards, quality of alternatives, and investments with satisfaction and commitment (Rusbult, 1980, Study 2); and (3) a multiwave longitudinal study (12 measurement occasions) in which participants completed a survey about their own relationship, assessing whether changes in satisfaction (costs and rewards), alternatives, and investments predicted subsequent commitment and relationship longevity (Rusbult, 1983).

Together these studies provided strong empirical evidence of several claims that were novel at the time and that launched a shift in relationship research from focusing exclusively on satisfaction to studying commitment processes more generally. The major claim was that satisfaction and commitment are not interchangeable, nor are they equally important in predicting relationship outcomes. Commitment was more strongly related to whether relationships endured than level of satisfaction (Rusbult, 1983). Understanding why some relationships persisted and others ended required understanding commitment, which increased with more rewards or higher satisfaction, with weakening alternatives, and with increasing investments. Whereas having more rewards consistently increased satisfaction, having greater costs associated with a relationship did not necessarily decrease satisfaction. Indeed, costs were not consistently related to commitment and even increased over time among those whose relationships endured (cf. Clark and Grote, 1998).

These initial tests of the investment model were major contributions to the study of relationships. In addition to providing a more complete and predictive account of enduring relationships, these initial tests accounted for findings that previously could not be explained. One such finding was that individuals left by their partner were very different than those who left their partner: Both decreased in their level of satisfaction, but those "abandoned" continued to invest heavily and had alternatives that declined in quality (Rusbult, 1983). That is, level of satisfaction could not differentiate the distinct processes that characterize "leavers" and the abandoned.

A second finding uniquely explained by the investment model was that rational individuals may persist in a relationship with an abusive partner. Victims of partner abuse experience low satisfaction, which might lead to the prediction that they would leave their partner. Prior to the investment model, it was widely believed that victims experiencing such negative events, and yet remaining with their partner, exhibit irrational, even pathological personal dispositions. In contrast, the investment model underscores structural features of the relationship that account for a victim remaining with an abusive partner: the victim may lack alternatives to the relationship and may have too much invested with that partner, making dissolution too costly. Indeed, Rusbult and Martz (1995) revealed that alternatives and investments were strongly related to whether battered women at a shelter remained committed and returned to their partner, whereas the association of satisfaction was weak or not significant depending on the measure of satisfaction.

More generally, these initial tests of the investment model launched a paradigm shift in the study of relationship processes (see Rusbult et al., 2006, and Agnew, 2009, for descriptions of other commitment models). The shift was from asking why people like [p. 221 ↓] each other, to asking how and why people stay together. Research following the development of the investment model identified specific processes by which committed individuals keep their relationship intact. It would be too easy and misguided to say, “They just want to stay together.” The model launched an entire area of research on various relationship maintenance phenomena that led to identifying thoughts and behaviors of committed individuals, and explaining the underlying processes that characterize these thoughts and behaviors. At the center of Rusbult’s theoretical account of relationship maintenance are the concepts of commitment and dependence, which we describe next.

Dependence, Commitment, and Relationship Maintenance

Dependence refers to the extent to which an individual “needs” a given relationship, or relies uniquely on that particular relationship for attaining desired outcomes. There are several processes through which individuals become dependent (Rusbult et al.,

1998). First, partners become dependent to the extent that they enjoy high satisfaction. Satisfaction level describes the degree to which an individual experiences positive versus negative affect as a result of involvement. Satisfaction level increases to the extent that a relationship gratifies the individual's most important needs, including needs for companionship, security, intimacy, sexuality, and belonging (Rusbult et al., 1998).

Dependence also increases when a person perceives that, on average, the best available alternative to a relationship is less desirable than the current relationship. Conversely, when a person's most important needs could be fulfilled outside of the current relationship – in a specific alternative relationship, by a combination of other involvements (by friends and family members, or on one's own) – a person's dependence on the current relationship diminishes. Interdependence theory argued that relationships would be more likely to endure when partners want to persist in a given relationship (i.e., satisfaction is high) and perceive they have no choice but to persist because they lack viable options to the relationship (i.e., alternatives are poor).

The investment model extended these claims in several respects. First, satisfaction and alternatives do not fully account for enduring relationships (Rusbult, 1980; Rusbult et al., 1998). If the decision to remain with or leave a partner were based solely on how positive one feels or on how one might anticipate feeling elsewhere, few relationships would endure – a relationship would collapse when positive feelings wane or when an attractive alternative becomes the target of one's attention. Relationships are not static, partners' affections ebb and flow, and many relationships persist in the face of tempting alternatives.

Second, dependence is influenced by high satisfaction, poor alternatives, and a third factor: investment size. Investment size refers to the magnitude and importance of the resources that become attached to a relationship that would be lost or decline in value if the relationship were to end. Partners form deep ties that bind themselves to each other by linking parts of themselves directly to the relationship – for example, investing their time and energy, disclosing personal information that ties their sense of dignity to the partner, sharing their own friends with the partner, and taking on shared possessions or giving things of value to the partner. Partners make such investments in the hope that doing so will create a strong foundation for a lasting future together. Investments increase dependence because the act of investment increases connections to the

partner that would be costly to break, in the same way that giving up a part of one's self is costly. As such, investments create a powerful psychological inducement to persist.

Third, the investment model extends prior theory by suggesting that commitment emerges as a consequence of increasing dependence (Rusbult et al., 1998).

Dependence is a [p. 222 ↓] structural property that describes the additive effects of satisfaction, investments, and (lack of) alternatives. When individuals want to persist (are satisfied), feel "tied into" the relationship or obliged to persist (have high investments), and have no choice but to persist (possess poor alternatives), they find themselves in circumstances objectively characterized as dependence.

As people become increasingly dependent they tend to develop strong commitment. Commitment level is defined as intent to persist in a relationship, including long-term orientation toward the involvement as well as feelings of psychological attachment to it (Arriaga and Agnew, 2001). When partners are satisfied, lack alternatives, and have invested heavily in their relationship, they form a strong intention to stay together, they see themselves as being connected (i.e., developing a strong relational identity and a sense of "we-ness;" Agnew et al., 1998), and adopt an orientation that reflects taking into account how things affect the long-term future of the relationship. As such, the psychological experience of commitment reflects more than the bases of dependence out of which it arises (i.e., high satisfaction, low alternatives, high investment). Commitment is the psychological state that directly influences everyday behavior in relationships and that mediates the effects of satisfaction, alternatives, and investments on behavior.

Having established that strong commitment – not high satisfaction – is the psychological state that characterizes partners in an enduring relationship, Rusbult, her colleagues, and many others have identified a multitude of commitment processes; that is, the many ways in which commitment promotes thoughts, feelings, and actions that, in turn, cause relationships to persist. Relationship maintenance is the upshot of responding to interpersonal situations by acting in the interest of the relationship. Past research has identified several relationship maintenance mechanisms through which highly committed people maintain their relationship.

Highly committed people are inclined to act in ways that promote relationship persistence. Their high commitment is particularly salient when they react to a challenging moment by doing what is best for the relationship. For example, when a partner makes a thoughtless remark or fails to follow through on a promise or acts in some other way that could damage the relationship, high commitment predicts accommodation, namely inhibiting the urge to retaliate and instead respond in ways that promote the relationship (Arriaga and Rusbult, 1998; Kilpatrick et al., 2002; Rusbult et al., 1991). Similarly, highly committed people are more inclined than their less committed counterparts to forego personal preferences for the sake of acting on behalf of the partner's interest (Powell and Van Vugt, 2003; Van Lange et al. 1997a, 1997b); and respond to a partner betrayal by forgiving the partner (Cann and Baucom, 2004; Finkel et al., 2002; McCullough et al., 1998). These relationship maintenance phenomena stem from strong commitment, not necessarily high satisfaction, a fact that would not be commonly accepted were it not for the investment model.

Highly committed people also think about things affecting the relationship differently than less committed people, and these thoughts make a difference in the wellbeing of a relationship. For example, committed people derogate tempting alternatives to shield against them, react to periods of doubt or uncertainty by denying negative qualities of the partner, develop unrealistically positive thoughts about their partner and/or the relationship, and cast others' relationships in a negative light (Agnew et al., 2001; Arriaga, 2002; Arriaga et al., 2007; Johnson and Rusbult, 1989; Lydon et al., 1999; Miller, 1997; Murray and Holmes, 1999; Murray et al., 1996; Rusbult et al., 2000; Simpson et al., 1990). Committed individuals also mentally see themselves in more relational terms – for example, they spontaneously use more plural pronouns – than less committed individuals (Agnew et al., 1998).

[p. 223 ↓]

Why do committed individuals come to think and act in a prorelationship manner? The interdependence theory distinction between the given situation and the effective situation provides some insight into this process (Kelley & Thibaut, 1978). The *given situation* refers to each partner's immediate, self-centered personal preferences in a specific situation. Within the context of a close relationship, of course, it is clear that people do not always pursue their given preferences. Behavior is often shaped

by broader concerns, including long-term goals to promote not only one's own but also one's partner's wellbeing. Movement away from given preferences results from *transformation of motivation*, a process which leads individuals to relinquish their immediate self-interest and act on the basis of broader considerations. The *effective situation* refers to the preferences resulting from the transformation process; effective preferences directly guide behaviors among those who are highly committed to their relationship.

The investment model premise that commitment and dependence are more consequential than satisfaction is the same premise that has guided research on mutual influence in deepening commitment and trust, or "mutual cyclical growth" (Agnew et al., 1998). As one person becomes more committed and acts in ways that reveal responsiveness to the partner and prorelationship tendencies, the partner becomes more comfortable being more dependent on and committed toward the person, which in turn makes the partner more likely to act in ways that reveal responsiveness and prorelationship tendencies, and so the process continues as each individual's commitment and trust in the other's responsiveness grows (Wieselquist et al., 1999). Although research on mutual cyclical growth does not directly examine all investment model components, it rests squarely on the idea that commitment, not satisfaction, brings about actions that will cause a relationship to persist.

Research on mutual cyclical growth underscores the manner in which relationship maintenance phenomena act to increase the partner's commitment. The partner essentially makes an attribution about the person acting in a prorelationship manner, and this attribution (that the person cares) fuels increases in the partner's investments, satisfaction, and commitment. Moreover, when a person acts in a prorelationship manner, such acts make salient to the person that he or she is committed and cares about the relationship unit (Agnew et al., 1998; Wieselquist et al., 1999). As such, relationship maintenance perceptions and acts likely influence couple members' satisfaction, alternatives, and investments and thus "effects" may become "causes." For example, over time, couple members who have personally sacrificed a lot for the partner will likely come to view such sacrifices as increases in their investments).

In short, the investment model has been extraordinarily generative in stimulating an entire research area that explains specifically how and why enduring relationships are

maintained, where others end. The investment model triggered a systematic analysis of much of what transpires in ongoing relationships. In the next section, we describe empirical tests assessing the generalizability of the investment model and review ways in which the model has stimulated research on a host of socially relevant topics.

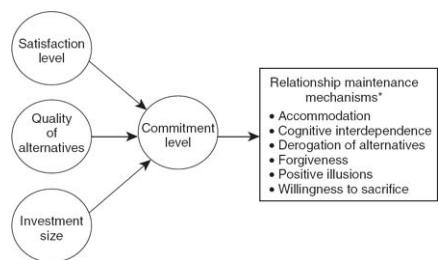
Generalizability and Empirical Robustness of the Investment Model

In the years since its initial testing, the investment model has been employed in a range of studies applying the model to participants of diverse ethnicities (Davis and Strube, 1993; Lin and Rusbult, 1995), homosexual and heterosexual partnerships (Duffy and Rusbult, 1986; Kurdek, 1991, 1995), abusive relationships (Choice and Lamke, 1999; Rhatigan and Axsom, 2006; Rusbult and Martz, 1995), socially marginalized relationships (Lehmiller and Agnew, 2006, 2007), and friendships [p. 224 ↓] (Hirofumi, 2003; Lin and Rusbult, 1995; Rusbult, 1980). In all of these studies, satisfaction level, quality of alternatives, and investment size are posited to have additive, main effects on commitment (see [Figure 37.1](#)). The model does not suggest that any one of the three predictors will be particularly influential in driving commitment. Rather, it suggests that all three factors may contribute to the prediction of commitment in an additive fashion. Multiple regression analyses have been used most often to test the model.

Although the majority of evidence supporting the investment model comes from studies of interpersonal relationships, the model also has been employed in other, nonrelational contexts (see Le and Agnew, 2003) with nonrelational targets of commitments. For instance, organizational and job commitment (cf. Farrell and Rusbult, 1981; Oliver, 1990) have been predicted in studies based on investment model constructs. In addition, Ping (1993, 1997) adapted the model to describe business interactions, and Lyons and Lowery (1989) have conceptualized commitment to one's residential community using a similar perspective. The investment model has been used successfully to predict patients' commitment to a medical regimen (Putnam et al., 1994), college students' commitment to their schools (cf. Geyer et al., 1987), and commitment to participating in musical activities (Koslowsky and Kluger, 1986). Finally, the sport commitment model has its roots firmly in the investment model (Raedeke, 1997;

Schmidt and Stein, 1991) and has been used to predict commitment of soccer and cricket players to their sports (Carpenter and Coleman, 1998; Carpenter and Scanlan, 1998).

Figure 37.1 The investment model of commitment processes



* Note that the specific relationship maintenance mechanisms listed here are based on extant empirical findings and that future research may uncover additional mechanisms.

As suggested above, the utility and robustness of the investment model has been demonstrated in numerous studies. A meta-analysis by Le and Agnew (2003) summarizes quantitative data regarding the model's performance, compiling empirical tests conducted through 1999. The meta-analysis included data from 52 studies (including 60 independent samples and over 11,000 participants). Overall, the average correlations between investment model constructs were found to be quite strong. Satisfaction level, quality of alternatives, and investment size each were highly correlated with commitment ($r = 0.68, -0.48, 0.46$ respectively), with the correlation between satisfaction and commitment found to be significantly stronger than the alternatives-commitment and investments-commitment correlations. The absolute magnitudes of the alternatives-commitment and [p. 225 ↓] investments-commitment correlations were not significantly different from one another. Satisfaction, alternatives, and investments were also found to be significantly correlated with one another (satisfaction-alternatives $r = -0.44$; satisfaction-investments $r = 0.42$; alternatives-investments $r = -0.25$).

Le and Agnew (2003) also examined the average standardized regression coefficients of commitment regressed simultaneously onto satisfaction, alternatives, and investments, and thus they assessed the relative independent contribution of each

variable in predicting commitment. Paralleling the correlational analyses, the meta-analysis revealed that satisfaction was the strongest predictor of commitment ($\text{std } \beta = 0.510$), whereas alternatives and investments were of similar absolute magnitude ($\text{std } \beta = -0.217$ and 0.240 respectively). In addition, 61 percent of the variance in commitment (95% CI [0.59, 0.63]) was accounted for by satisfaction, alternatives, and investments collectively. Moreover, moderator analyses suggested that the associations between commitment and its theorized bases vary minimally as a function of demographic (e.g., ethnicity) or relational (e.g., duration) factors.

Commitment is held to mediate the effects of satisfaction, alternatives, and investments on consequential behaviors. The meta-analysis also provided support for this supposition. Specifically, the correlation between commitment and later stay-leave behaviors reported by participants (e.g., whether the couple was still together or the worker was still at the job) across 12 studies was found to be 0.47 (95% CI [0.43, 0.50], $N = 1720$).

Recent Applications of the Investment Model

Since 1999 (the inclusion year cut-off in the Le and Agnew meta-analysis), dozens of additional studies have been published that test the investment model or aspects of it. Some of these papers confirm the findings from earlier publications on the applicability of the investment model in understanding commitment in various types of relationships, beyond romantic involvements, such as instructors' commitment to student supervision (Peleg-Oren et al., 2007), or parental commitment to their child's pediatrician (Agnew and VanderDrift, 2010).

Recent research also has confirmed findings from earlier publications on the applicability of the investment model in understanding commitment to nonperson targets. For example, the investment model provides predictive value in understanding employees' attitudes toward different job changes (e.g., changing department or relocating to a different office; van Dam, 2005), clients' commitment to their bank (Kastlunger et al., 2008), childcare providers' commitment to the childcare center that

employs them (Gable and Hunting, 2001), and customer loyalty to specific brands (Li and Petrick, 2008).

A recent meta-analysis of predictors of nonmarital romantic relationship dissolution also provides additional support for the claim that commitment is a key proximal predictor of stay–leave behavior. Using data collected from nearly 38,000 participants participating in 137 studies over a 33-year period, commitment was found to be a particularly powerful predictor of breakup (Le et al., 2010). Specifically, assuming random effects (Lipsey and Wilson, 2001), the weighted mean effect size (d) for commitment in predicting premarital breakup was found to be -0.832 (95% CI [-0.934, -0.729]).

Recently, the investment model has also been applied to understanding public support for government actions regarding foreign policies. Existing public opinion research often focuses on factors such as partisanship to understand where people stand on foreign policy issues (e.g., Zaller, 1994). By contrast, the investment model would predict that people use specific policy performance criteria to determine the value of persisting with the same policy. Agnew and colleagues (Agnew et al., 2007) used the investment model to examine commitment to the “war on terror” waged by the US under President [p. 226 ↓] George W. Bush. They conducted two experiments in which they simultaneously manipulated the three investment model constructs (a novel aspect of this research). As predicted by the investment model, participants were most strongly committed to the war on terror when satisfaction with reported outcomes of the war (e.g., reducing threat to the citizens of the US) and investments in the war (e.g., casualties) were described as high and alternatives (e.g., diplomatic solutions) were described as low. In contrast, participants were least committed to the war when satisfaction and investments were described as low and alternatives were described as high. Similarly, Hoffman et al. (2009) found significant support for the investment model in predicting citizen commitment to their country being a part of both the United Nations and NATO.

Other researchers have used the investment model as a springboard for examining related issues pertinent to commitment. Such papers have not reported formal tests of the overall model; rather, they have used it as the basis for furthering understanding of variables or processes beyond those specified in the investment model. For example, in an analysis of how narcissism relates to commitment, Campbell and Foster (2002)

found the association between narcissism and commitment is negative and largely mediated by perceptions of alternatives. Katz et al. (2006) found that women involved in sexually coercive relationships reported greater investment in their relationship but did not differ from other women in satisfaction or commitment. Vaculík and Jedrzejczyková (2009) focused on describing differences between people involved in various types of unmarried cohabitation and used investment model variables to characterize such differences. Taking a social cognitive approach, Etcheverry and Le (2005) tested and found support for the notion that the cognitive accessibility of commitment moderates the association between self-reported commitment and relationship persistence, accommodative responses, and willingness to sacrifice. Finally, in their examination of safer-sex behavior among committed gay male partners, Davidovich et al. (2006) found that low satisfaction with the relationship was associated with more risky unprotected anal intercourse, whereas high commitment to the relationship was associated with greater efforts toward the practice of safer sex. All of these studies speak to the generalizability of the investment model beyond its originally envisioned use.

Extending the Investment Model

In this last section, we describe a recent extension to the investment model that attempts to account for additional variance in levels of commitment. We begin by focusing on investments. Goodfriend and Agnew (2008) have elaborated on the investment concept, suggesting that the notion of investments should include not only things that have already been invested, but also any *plans* that partners have made, either individually or with the partner, regarding the relationship. In ending a relationship, one loses not only those investments that have been sunken to date but also the possibility of achieving any future plans with the partner. Thus, the plans that one forms with a partner act to keep one's commitment to the partnership alive. One notable aspect of considering future plans as contributing to current commitment is that such plans do not require that relationship partners have much of a shared history together. That is, even partners who have known one another for a relatively short period of time may become quite committed to continuing their relationship, not because of considerable past sunken costs but because of a motivation to see cherished future plans come to fruition. Goodfriend and Agnew (2008) found that future plans were

strongly predictive of romantic relationship commitment, above and beyond past investments.

Recent research has also examined how others outside of the dyad might influence a [p. 227 ↓] couple's motivation to continue a relationship. Research has examined the associations between perceptions of a social network's approval or disapproval for a romantic relationship and characteristics of that relationship (Agnew et al., 2001; Arriaga et al., 2004; Bryant and Conger, 1999; Cox et al., 1997; Lehmiller and Agnew, 2006, 2007; Loving, 2006; Parks et al., 1983; Sprecher 1988; Sprecher and Felmlee, 1992). In general, past studies have shown that qualities, structure, and opinions of social network members are associated with the quality and functioning of dyadic relationships embedded in that network.

Research that integrates the investment model with research on social networks has examined the perceived role of social referents: A couple's commitment can be affected by their perceptions of what important others think of their relationship, as well as the couple members' motivation to follow what others think (i.e., the couple members' "subjective norms;" Ajzen and Fishbein, 1980; Fishbein and Ajzen, 1975). Etcheverry and Agnew (2004) found that subjective norms provided additional prediction of relationship commitment, above and beyond the effects of satisfaction, alternatives, and past investments. Moreover, just as behavioral intention mediates the effect of subjective norms on behavior in the theory of reasoned action (Ajzen and Fishbein, 1980), longitudinal analyses indicated that commitment mediated the effect of subjective norms on remaining in a romance approximately 8 months later. Both theoretically and empirically, the subjective norms construct broadens the prediction of relationship commitment beyond the original three predictors of the investment model.

Combining these recent theoretical and empirical advances, Agnew et al. (2008) proposed a new model to account not only for continuity in close relationships but also for possible changes in type of relationship with a given partner (e.g., shifts from romantic involvement to friendship, or vice versa). They proposed the bases of relational commitment model (BORC model), which mirrors the investment model in specifying three predictors of commitment. One predictor combines satisfaction and alternatives by focusing on outcomes relative to standards, harking back to some of the early interdependence writing on outcomes in dyadic interaction (Thibaut and

Kelley, 1959). A second predictor is an expanded conceptualization of investments – what they refer to as *valued linkages* – which incorporates recent work on future plans (Goodfriend and Agnew, 2008). A third predictor, subjective norms, incorporates work by Etcheverry and Agnew (2004) showing that couple members' commitment is affected by what important others think about their relationship. Initial tests of the BORC model have yielded supportive findings, with each of the model's specified variables accounting for unique variance in relationship type commitment and the overall model accounting for over three-quarters of the variance in commitment. Of course, the foundation of this model is “standing on the shoulder of giants,” benefiting from the decades of accumulated knowledge about commitment processes originally set into motion by the investment model.

Conclusion

Most social psychologists are familiar with Kurt Lewin's famous statement, “There is nothing so practical as a good theory” (1951: 169). Lewin would have loved the investment model. The model has provided an extremely practical theoretical framework for understanding and explaining the causes and consequences of commitment. It originated as a model to understand why people remain in romantic relationships. Subsequently, it has been used to examine commitment in all kinds of relationship and to all types of targets. It has also been utilized to examine the specific ways in which commitment brings about persistence, the specific thoughts and actions that differentiate people based on their level of commitment.

[p. 228 ↓]

There are many ways in which the investment model has had an impact not just on the study of close relationships and the field of social psychology, but also more broadly. First, this model transformed the way scholars in various fields think about commitment. Without the investment model, our understanding of various commitment processes would be diminished – relationship commitment, yes, but also organizational commitment, sports commitment, policy commitment, and so on. Second, the investment model provided a critical new direction in the study of relationships. This new direction comprised a social psychological analysis, vis-à-vis a structural

sociological analysis, and it focused on the *ongoing* course of intimate relationships, rather than focusing on their onset alone. Third, by explicitly modeling a psychological process, the investment model advanced a *scientific analysis* of relationships. As we move into the future, we hope and expect that the investment model will continue to advance theory and applications in social psychology and beyond.

Notes

1 With the exception of the section on origins of the theory, this chapter was prepared following the death of the first author, Caryl Rusbult, our mentor and dear friend.

References

Agnew, C.R. (2009) Commitment, theories and typologies . In H.T. Reis, ed. and S.K. Sprecher (eds), Encyclopedia of Human Relationships , 1, 245–248. Thousand Oaks, CA: Sage. <http://dx.doi.org/10.4135/9781412958479>

Agnew, C.R., Arriaga, X.B. and Wilson, J.E. (2008) Committed to what? Using the Bases of Relational Commitment Model to understand continuity and changes in social relationships . In J.P. Forgas, ed. and J. Fitness (eds), Social Relationships: Cognitive, Affective and Motivational Processes , pp. 147–164. New York: Psychology Press.

Agnew, C.R., Hoffman, A.M., Lehmiller, J.J. and Duncan, N.T. From the interpersonal to the international: Understanding commitment to the ‘War on Terror’ . Personality and Social Psychology Bulletin , 33, 1559–1571. (2007) <http://dx.doi.org/10.1177/0146167207305540>

Agnew, C.R., Loving, T.J. and Drigotas, S.M. Substituting the forest for the trees: Social networks and the prediction of romantic relationship state and fate . Journal of Personality and Social Psychology , 81, 1042–1057. (2001) <http://dx.doi.org/10.1037/0022-3514.81.6.1042>

Agnew, C.R. and VanderDrift, L.E. (2010) Commitment within the patient-physician relationship: Parental commitment to pediatricians .

Agnew, C.R., Van Lange, P.A.M., Rusbult, C.E. and Langston, C.A. Cognitive interdependence: Commitment and the mental representation of close relationships . Journal of Personality and Social Psychology , 74, 939–954. (1998) <http://dx.doi.org/10.1037/0022-3514.74.4.939>

Ajzen, I. and Fishbein, M. (1980) Understanding Attitudes and Predicting Social Behavior . Englewood Cliffs, NJ: Prentice Hall.

Arriaga, X.B. Joking violence among highly committed individuals . Journal of Interpersonal Violence , 17, 591–610. (2002) <http://dx.doi.org/10.1177/0886260502017006001>

Arriaga, X.B. and Agnew, C.R. Being committed: Affective, cognitive, and conative components of relationship commitment . Personality and Social Psychology Bulletin , 27, 1190–1203. (2001) <http://dx.doi.org/10.1177/0146167201279011>

Arriaga, X.B., Goodfriend, W. and Lohmann, A. (2004) Beyond the individual: Concomitants of closeness in the social and physical environment . In D. Mashek, ed. and A. Aron (eds), Handbook on Relationship Closeness , pp. 287–303. Mahwah, NJ: Erlbaum.

Arriaga, X.B. and Rusbult, C.E. Standing in my partner's shoes: Partner perspective-taking and reactions to accommodative dilemmas . Personality and Social Psychology Bulletin , 24, 927–948. (1998) <http://dx.doi.org/10.1177/0146167298249002>

Arriaga, X.B., Slaughterbeck, E.S., Capezza, N.M. and Hmurovic, J.L. From bad to worse: Relationship commitment and vulnerability to partner imperfections . Personal Relationships , 14, 389–409. (2007) <http://dx.doi.org/10.1111/j.1475-6811.2007.00162.x>

Becker, H.S. Notes on the concept of commitment . American Journal of Sociology , 66, 32–40. (1960) <http://dx.doi.org/10.1086/222820>

Blau, P.M. (1967) Exchange and Power in Social Life . New York: Wiley.

Brockner, J., Shaw, M.C. and Rubin, J.Z. Factors affecting withdrawal from an escalating conflict: Quitting before it's too late . *Journal of Experimental Social Psychology* , 15, 492–503. (1979) <http://dx.doi.org/10.1016/0022-1031%2879%2990011-8>

Bryant, C.M. and Conger, R.D. Marital success and domains of social support in long-term relationships: Does the influence of network members ever end? *Journal of Marriage and the Family* , 61, 437–450. (1999) <http://dx.doi.org/10.2307/353760>

Campbell, W.K. and Foster, C.A. Narcissism and commitment in romantic relationships: An investment model analysis . *Personality and Social Psychology Bulletin* , 28, 484–495. (2002). <http://dx.doi.org/10.1177/0146167202287006>

Cann, A. and Baucom, T.R. Former partners and new rivals as threats to a relationship: Infidelity type, gender, and commitment as factors related to distress and forgiveness . *Personal Relationships* , 11, 305–318. (2004) <http://dx.doi.org/10.1111/j.1475-6811.2004.00084.x>

Carpenter, P.J. and Coleman, R. A longitudinal test of elite youth cricketers' commitment . *International Journal of Sport Psychology* , 29, 195–210. (1998).

Carpenter, P.J. and Scanlan, T.K. Changes over time in determinants of sport commitment . *Pediatric Exercise Science* , 10, 356–365. (1998)

Choice, P. and Lamke, L. Stay/leave decision-making in abusive dating relationships . *Personal Relationships* , 6, 351–368. (1999) <http://dx.doi.org/10.1111/j.1475-6811.1999.tb00197.x>

Clark, M.S. and Grote, N.K. Why aren't indices of relationship costs always negatively related to indices of relationship quality? *Personality and Social Psychology Review* , 2, 2–17. (1998) http://dx.doi.org/10.1207/s15327957pspr0201_1

Cox, C.L., Wexler, M.O., Rusbult, C.E. and Gaines, S.O.Jr. Prescriptive support and commitment processes in close relationships . *Social Psychology Quarterly* , 60, 79–90. (1997) <http://dx.doi.org/10.2307/2787013>

Davidovich, U., de Wit, J. and Stroebe, W. Relationship characteristics and risk of HIV Infection: Rusbult's investment model and sexual risk behavior of gay men in steady relationships . *Journal of Applied Social Psychology* , 36, 22–40. (2006) <http://dx.doi.org/10.1111/j.0021-9029.2006.00002.x>

Davis, L. and Strube, M.J. An assessment of romantic commitment among black and white dating couples . *Journal of Applied Social Psychology* , 23, 212–225. (1993) <http://dx.doi.org/10.1111/j.1559-1816.1993.tb01083.x>

Duffy, S.M. and Rusbult, C.E. Satisfaction and commitment in homosexual and heterosexual relationships . *Journal of Homosexuality* , 12, 1–23. (1986) http://dx.doi.org/10.1300/J082v12n02_01

Etcheverry, P.E. and Agnew, C.R. Subjective norms and the prediction of romantic relationship state and fate . *Personal Relationships* , 11, 409–428. (2004) <http://dx.doi.org/10.1111/j.1475-6811.2004.00090.x>

Etcheverry, P.E. and Le, B. Thinking about commitment: Accessibility of commitment and prediction of relationship persistence, accommodation, and willingness to sacrifice . *Personal Relationships* , 12, 103–123. (2005) <http://dx.doi.org/10.1111/j.1350-4126.2005.00104.x>

Farrell, D. and Rusbult, C.E. Exchange variables as predictors of job satisfaction, job commitment, and turnover: The impact of rewards, costs, alternatives, and investments . *Organizational Behavior and Human Performance* , 28, 78–95. (1981) <http://dx.doi.org/10.1016/0030-5073%2881%2990016-7>

Finkel, E.J., Rusbult, C.E., Kumashiro, M. and Hannon, P.A. Dealing with betrayal in close relationships: Does commitment promote forgiveness? *Journal of Personality and Social Psychology* , 82, 956–974. (2002) <http://dx.doi.org/10.1037/0022-3514.82.6.956>

Fishbein, M. and Ajzen, I. (1975) *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research* . Reading, MA: Addison-Wesley.

Gable, S. and Hunting, M. Child care providers' organizational commitment: A test of the investment model . *Child and Youth Care Forum* , 30, 265–281. (2001) <http://dx.doi.org/10.1023/A:1014425310598>

Geyer, P.D., Brannon, Y.S. and Shearon, R.W. The prediction of students' satisfaction with community college vocational training . *The Journal of Psychology* , 121, 591–597. (1987) <http://dx.doi.org/10.1080/00223980.1987.9712688>

Goodfriend, W. and Agnew, C.R. Sunken costs and desired plans: Examining different types of investments in close relationships . *Personality and Social Psychology Bulletin* , 34, 1639–1652. (2008) <http://dx.doi.org/10.1177/0146167208323743>

Hirofumi, A. Closeness and interpersonal outcomes in same-sex friendships: An improvement of the investment model and explanation of closeness . *Japanese Journal of Experimental Social Psychology* , 42, 131–145. (2003) <http://dx.doi.org/10.2130/jjesp.42.131>

Hoffman, A.M., Agnew, C.R., Lehmiller, J.J. and Duncan, N.T. Satisfaction, alternatives, investments, and the microfoundations of audience cost models . *International Interactions* , 35, 365–389. (2009) <http://dx.doi.org/10.1080/03050620903328274>

Johnson, D.J. and Rusbult, C.E. Resisting temptation: Devaluation of alternative partners as a means of maintaining commitment in close relationships . *Journal of Personality and Social Psychology* , 57, 967–980. (1989) <http://dx.doi.org/10.1037/0022-3514.57.6.967>

Kastlunger, B., Martini, M., Kirchler, E. and Hofmann, E. Impegno, soddisfazione e fiducia del cliente bancario: Un'analisi empirica a Roma e in Sardegna . *Psicologia Sociale* , 3, 307–324. (2008)

Katz, J., Kuffel, S.W. and Brown, F.A. Leaving a sexually coercive dating partner: A prospective application of the investment model . *Psychology of Women Quarterly* , 30, 267–275. (2006) <http://dx.doi.org/10.1111/j.1471-6402.2006.00295.x>

Kelley, H.H. and Thibaut, J.W. (1978) *Interpersonal Relations: A Theory of Interdependence* . New York: Wiley.

Kiesler, C.A. (1971) *The Psychology of Commitment: Experiments Linking Behavior to Belief*. New York: Academic Press.

Kilpatrick, S.D., Bissonnette, V.L. and Rusbult, C.E. Empathic accuracy and accommodative behavior among newly married couples . *Personal Relationships* , 9, 369–393. (2002) <http://dx.doi.org/10.1111/1475-6811.09402>

Koslowsky, M. and Kluger, A. Commitment to participation in musical activities: An extension and application of the investment model . *Journal of Applied Social Psychology* , 16, 831–844. (1986) <http://dx.doi.org/10.1111/j.1559-1816.1986.tb01171.x>

Kurdek, L.A. Correlates of relationship satisfaction in cohabiting gay and lesbian couples: Integration of contextual, investment, and problem-solving models . *Journal of Personality and Social Psychology* , 61, 910–922. (1991) <http://dx.doi.org/10.1037/0022-3514.61.6.910>

Kurdek, L.A. Assessing multiple determinants of relationship commitment in cohabiting gay, cohabiting lesbian, dating heterosexual, and married heterosexual couples . *Family Relations* , 44, 261–266. (1995) <http://dx.doi.org/10.2307/585524>

Le, B. and Agnew, C.R. Commitment and its theorized determinants: A meta-analysis of the investment model . *Personal Relationships* , 10, 37–57. (2003) <http://dx.doi.org/10.1111/1475-6811.00035>

Le, B., Dove, N., Agnew, C.R., Korn, M.S. and Mutso, A.A. Predicting non-marital romantic relationship dissolution: A meta-analytic synthesis . *Personal Relationships* , 17, 377–390. (2010) <http://dx.doi.org/10.1111/j.1475-6811.2010.01285.x>

Lehmiller, J.J. and Agnew, C.R. Marginalized relationships: The impact of social disapproval on romantic relationship commitment . *Personality and Social Psychology Bulletin* , 32, 40–51. (2006) <http://dx.doi.org/10.1177/0146167205278710>

Lehmiller, J.J. and Agnew, C.R. Perceived marginalization and the prediction of romantic relationship stability . *Journal of Marriage and Family* , 69, 1036–1049. (2007). <http://dx.doi.org/10.1111/j.1741-3737.2007.00429.x>

Lewin, K. (1951) Field Theory in Social Science; Selected Theoretical Papers . In D. Cartwright (ed.). New York: Harper & Row.

Li, X. and Petrick, J.F. Examining the antecedents of brand loyalty from an investment model perspective . *Journal of Travel Research* , 47, 25–34. (2008) <http://dx.doi.org/10.1177/0047287507312409>

Lin, Y.W. and Rusbult, C.E. Commitment to dating relationships and cross-sex friendships in America and China . *Journal of Social and Personal Relationships* , 12, 7–26. (1995) <http://dx.doi.org/10.1177/0265407595121002>

Lipsey, M.W. and Wilson, D.B. (2001) Practical Meta-analysis . Thousand Oaks, CA: Sage Publications.

Loving, T.J. Predicting dating relationship fate with insiders' and outsiders' perspectives: Who and what is asked matters . *Personal Relationships* , 13, 349–362. (2006) <http://dx.doi.org/10.1111/j.1475-6811.2006.00122.x>

Lydon, J., Meana, M., Sepinwall, D., Richards, N. and Mayman, S. The commitment calibration hypothesis: When do people devalue attractive alternatives? *Personality and Social Psychology Bulletin* , 25, 152–161. (1999) <http://dx.doi.org/10.1177/0146167299025002002>

Lyons, W.E. and Lowery, D. Citizen responses to dissatisfaction in urban communities: A partial test of a general model . *Journal of Politics* , 51, 841–868. (1989) <http://dx.doi.org/10.2307/2131537>

McCullough, M.E., Rachal, K.C., Sandage, S.J., Worthington, E.L., Jr., Brown, S.W. and Hight, T.L. Interpersonal forgiving in close relationships II: Theoretical elaboration and measurement . *Journal of Personality and Social Psychology* , 75, 1586–1603. (1998) <http://dx.doi.org/10.1037/0022-3514.75.6.1586>

Miller, R.S. Inattentive and contented: Relationships commitment and attention to alternatives . *Journal of Personality and Social Psychology* , 73, 758–766. (1997) <http://dx.doi.org/10.1037/0022-3514.73.4.758>

Murray, S.L. and Holmes, J.G. The (mental) ties that bind: Cognitive structures that predict relationship resilience . *Journal of Personality and Social Psychology* , 77, 1228–1244. (1999) <http://dx.doi.org/10.1037/0022-3514.77.6.1228>

Murray, S.L., Holmes, J.G. and Griffin, D.W. The self-fulfilling nature of positive illusions in romantic relationships: Love is not blind, but prescient . *Journal of Personality and Social Psychology* , 71, 1155–1180. (1996) <http://dx.doi.org/10.1037/0022-3514.71.6.1155>

Oliver, N. Rewards, investments, alternatives and organizational commitment: Empirical and evidence and theoretical development . *Journal of Occupational Psychology* , 63, 19–31. (1990) <http://dx.doi.org/10.1111/j.2044-8325.1990.tb00507.x>

Parks, M.R., Stan, C.M. and Eggert, L.L. Romantic involvement and social network involvement . *Social Psychology Quarterly* , 46, 116–131. (1983) <http://dx.doi.org/10.2307/3033848>

Peleg-Oren, N., Macgowan, M.J. and Even-Zahav, R. Field instructors' commitment to Student Supervision: Testing the investment model . *Social Work Education* , 26, 684–696. (2007) <http://dx.doi.org/10.1080/02615470601129875>

Ping, R.A., Jr. The effects of satisfaction and structural constraints on retailer exiting, voice, loyalty, opportunism, and neglect . *Journal of Retailing* , 69, 320–352. (1993) <http://dx.doi.org/10.1016/0022-4359%2893%2990010-G>

Ping, R.A., Jr. Voice in business-to-business relationships: Cost-of-exit and demographic antecedents . *Journal of Retailing* , 73, 261–281. (1997) <http://dx.doi.org/10.1016/S0022-4359%2897%2990006-3>

Powell, C. and Van Vugt, M. Genuine giving or selfish sacrifice?: The role of commitment and cost level upon willingness to sacrifice . *European Journal of Social Psychology* , 33, 403–412. (2003) <http://dx.doi.org/10.1002/ejsp.154>

Putnam, D.E., Finney, J.W., Barkley, P.L. and Bonner, M.J. Enhancing commitment improves adherence to a medical regimen . *Journal of Consulting and Clinical Psychology* , 62, 191–194. (1994) <http://dx.doi.org/10.1037/0022-006X.62.1.191>

Raedeke, T.D. Is athlete burnout more than just stress? A sport commitment perspective . *Journal of Sport and Exercise Psychology* , 19, 396–417. (1997)

Rhatigan, D.L. and Axsom, D.K. Using the investment model to understand battered women's commitment to abusive relationships . *Journal of Family Violence* , 21, 153–162. (2006) <http://dx.doi.org/10.1007/s10896-005-9013-z>

Rusbult, C.E. Commitment and satisfaction in romantic associations: A test of the investment model . *Journal of Experimental Social Psychology* , 16, 172–186. (1980) <http://dx.doi.org/10.1016/0022-1031%2880%2990007-4>

Rusbult, C.E. A longitudinal test of the investment model: The development (and deterioration) of satisfaction and commitment in heterosexual involvements . *Journal of Personality and Social Psychology* , 45, 101–117. (1983) <http://dx.doi.org/10.1037/0022-3514.45.1.101>

Rusbult, C.E., Coolsen, M.K., Kirchner, J.L. and Clarke, J.A. (2006) Commitment . In A.L. Vangelisti, ed. and D. Perlman (eds), *The Cambridge Handbook of Personal Relationships* , pp. 615–635. New York: Cambridge University Press.

Rusbult, C.E. and Martz, J.M. Remaining in an abusive relationship: An investment model analysis of nonvoluntary dependence . *Personality and Social Psychology Bulletin* , 21, 558–571. (1995) <http://dx.doi.org/10.1177/0146167295216002>

Rusbult, C.E., Martz, J.M. and Agnew, C.R. The investment model scale: Measuring commitment level, satisfaction level, quality of alternatives, and investment size . *Personal Relationships* , 5, 357–391. (1998) <http://dx.doi.org/10.1111/j.1475-6811.1998.tb00177.x>

Rusbult, C.E., Van Lange, P.A.M., Wildschut, T., Yovetich, N.A. and Verette, J. Perceived superiority in close relationships: Why it exists and persists . *Journal of Personality and Social Psychology* , 79, 521–545. (2000) <http://dx.doi.org/10.1037/0022-3514.79.4.521>

Rusbult, C.E., Verette, J., Whitney, G.A., Slovik, L.F. and Lipkus, I. Accommodation processes in close relationships: Theory and preliminary empirical evidence .

Journal of Personality and Social Psychology , 60, 53–78. (1991) <http://dx.doi.org/10.1037/0022-3514.60.1.53>

Schelling, J. An essay on bargaining . American Economic Review , 46, 281–306. (1956)

Schmidt, G.W. and Stein, G.L. Sport commitment: A model integrating enjoyment, dropout, and burnout . Journal of Sport and Exercise Psychology , 13, 254–265. (1991)

Shubik, M. The dollar auction game: A paradox in noncooperative behavior and escalation . Journal of Conflict Resolution , 15, 109–111. (1971) <http://dx.doi.org/10.1177/002200277101500111>

Simpson, J.A., Gangestad, S.W. and Lerma, M. Perception of physical attractiveness: Mechanisms involved in the maintenance of romantic relationships . Journal of Personality and Social Psychology , 59, 1192–1201. (1990) <http://dx.doi.org/10.1037/0022-3514.59.6.1192>

Sprecher, S. Investment model, equity, and social support determinants of relationship commitment . Social Psychology Quarterly , 51, 318–328. (1988) <http://dx.doi.org/10.2307/2786759>

Sprecher, S. and Felmlee, D. The influence of parents and friends on the quality and stability of romantic relationships: A three-wave longitudinal investigation . Journal of Marriage and the Family , 54, 888–900. (1992) <http://dx.doi.org/10.2307/353170>

Staw, B.M. Knee-deep in the big muddy: A study of escalating commitment to a chosen action . Organi zation Behavior and Human Performance , 16, 27–44. (1976) <http://dx.doi.org/10.1016/0030-5073%2876%2990005-2>

Teger, A.I. (1980) Too Much Invested to Quit . New York: Pergamon.

Thibaut, J.W. and Kelley, H.H. (1959) The Social Psychology of Groups . New York: Wiley.

Tropper, R. The consequences of investment in the process of conflict . Journal of Conflict Resolution , 16, 97–98. (1972) <http://dx.doi.org/10.1177/002200277201600107>

Vaculík, M. and Jedrzejczyková, V. Commitment in unmarried cohabitation . Studia Psychologica , 51, 101–117. (2009)

van Dam, K. Employee attitudes toward job changes: An application and extension of Rusbult and Farrell's investment model . Journal of Occupational and Organizational Psychology , 78, 253–272. (2005) <http://dx.doi.org/10.1348/096317904X23745>

Van Lange, P.A.M., Agnew, C.R., Harnick, F. and Steemers, G.E.M. From game theory to real life: How social value orientation affects willingness to sacrifice in ongoing close relationships . Journal of Personality and Social Psychology , 73, 1330–1344. (1997a) <http://dx.doi.org/10.1037/0022-3514.73.6.1330>

Van Lange, P.A.M., Rusbult, C.E., Drigotas, S.M., Arriaga, X.B., Witcher, B.S. and Cox, C.L. Willingness to sacrifice in close relationships . Journal of Personality and Social Psychology , 72, 1373–1395. (1997b) <http://dx.doi.org/10.1037/0022-3514.72.6.1373>

Wieselquist, J., Rusbult, C.E., Foster, C.A. and Agnew, C.R. Commitment, pro-relationship behavior, and trust in close relationships . Journal of Personality and Social Psychology , 77, 942–966. (1999) <http://dx.doi.org/10.1037/0022-3514.77.5.942>

Zaller, J. (1994) Strategic politicians, public opinion, and the Gulf crisis . In W.L. Bennett, ed. and D.L. Paletz (eds), Taken by Storm: The Media, Public Opinion, and U.S. Foreign Policy in the Gulf War , pp. 250–276. Chicago: University of Chicago Press.

<http://dx.doi.org/10.4135/9781446249222.n37>

Handbook of Theories of Social Psychology

A Theory of Communal (and Exchange) Relationships

Contributors: Margaret S. Clark & Judson R. Mills

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "A Theory of Communal (and Exchange) Relationships"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n38>

Print pages: 232-251

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n38>

[p. 232 ↓]

Chapter 38: A Theory of Communal (and Exchange) Relationships

Margaret S.Clark¹ Judson R.Mills

Abstract

A qualitative distinction between communal and exchange relationships (Clark and Mills, 1979; Mills and Clark, 1982) together with a quantitative dimension to communal relationships (Mills and Clark, 1982; Mills et al., 2004) is described. A review is given of empirical work supporting the theory and its implications for such things as: non-contingent helping; donor and recipient reactions to giving help; emotional expression and reactions to others' emotional expressions; keeping track of contributions and of needs in relationships; and, more broadly, what constitutes healthy and unhealthy intimate and non-intimate relationships. In the process, the theory, details of its development, refinement, and testing as well as challenges to our approach from other researchers are commented upon and placed in historical perspective.

Introduction

Why, when purchasing a gift for a friend, do we expect price tags to be on items yet, after the purchase, we make sure they been removed? Why did a friend who rented a vacation house and arrived to find no hot water, find it maddening the real estate agent tried to elicit her sympathy by explaining that the owners were experiencing severe personal problems? When we began work on a distinction between communal and exchange relationships in the 1970s, social psychologists did not have ready answers

to these questions. Yet the questions were intriguing and we set to work to provide theoretical answers backed with empirical support.

At that time, the now flourishing area of social research on close relationships had yet to emerge. There was work on interpersonal attraction to be sure. There was work on norms governing how people regulated the giving and receiving of benefits and rewards in relationships with equity theory being the most prominent (see, for example, Adams, 1965; Messick and Cook, 1983; Walster et al., 1978). Yet no social psychologist had suggested the possibility that the rules governing behavior might differ by relationship context. It was in that atmosphere that we set forth a qualitative distinction between communal and exchange relationships (Clark and Mills, 1979).

[p. 233 ↓]

We drew our inspiration from some brief observations made by sociologist Irving Goffman in his book *The Presentation of Self in Everyday Life*. Goffman had noted differences in the nature “social” and “economic” exchange. Social exchange, he said, was compromised by agreement in advance as to what was to be exchanged. Something given in a social exchange “need only be returned if the relationship calls for it; that is when the putative recipient comes to be in need of a favor or when he is ritually stationed for a ceremonial expression of regard.” In contrast, in economic exchange, “no amount of mere thanks can presumably satisfy the giver; he must get something of equivalent material value in return.” (Goffman, 1961: 275–276). The distinction was made briefly, no experimentation nor indeed any systematic research had been done to document it, and we did not completely agree with Goffman. Yet we thought his comments were important.

A Qualitative Distinction between Exchange and Communal Relationships

The qualitative distinction we initially drew between exchange and communal relationships dealt with the norms that governed the giving and acceptance of benefits. We defined the term benefit as something one member of a relationship chooses

to give to the other that is, in the donor's opinion (and typically in the recipient's and outside observers' as well), of use or value. Benefits take many forms. Services, goods, compliments, provision of information, supporting a person in reaching a goal, and symbols of caring such as cards or flowers, can all be benefits. Importantly, benefits are not the same as rewards. For our purposes the term "reward" refers to all pleasures, satisfactions or gratifications that the recipient might enjoy (Thibaut and Kelley, 1959: 12). For instance, a person might enjoy the reflected glory of being associated with a famous individual but unless the famous person chose to associate with the person *in order* to confer that reflected glory the enjoyment is a reward, not a benefit. Also, not all benefits constitute rewards. A person may give another person a benefit intending to meet the person's needs or desires but the benefit might *not* be a reward from the recipient's perspective. One of our students once received a bouquet from an admirer. The flowers were a benefit. They had value and the donor intended to give them to the recipient. The recipient generally liked flowers, yet she did not experience receiving *those* flowers as rewarding due to her lack of any desire for a relationship with the donor.²

The Nature of Exchange Relationships

In our initial papers (Clark and Mills, 1979; Mills and Clark, 1982; Clark, 1985) we posited that in many relationships members assume that a benefit is given with the expectation of receiving a comparable benefit (or benefits) in return. We chose the term "exchange relationship" as a label for these relationships rather than Goffman's term "economic exchange" because many of the benefits people give and receive do not involve money or things for which a monetary value can easily be calculated. Yet such benefits can still be exchanged, one for another. In these relationships, we said, the receipt of a benefit incurs an obligation (debt) to return a comparable benefit. In such relationships each person is concerned with how much he or she will receive in exchange for benefiting the other and how much is owed for benefits received. (We were using, and still use, the term "exchange" in a narrower sense than some others in the fields of social psychology and sociology.) Exchange relationships are often (but not always) exemplified by relationships that are called, in lay language, business relationships, relationships between acquaintances, and relationships between

strangers meeting and interacting [p. 234 ↓] for the first time (assuming no desire for a friendship or romantic relationship.)

The Nature of Communal Relationships

Not all relationships follow exchange rules. In some relationships benefits are given in support of the partner's welfare *non-contingently*; that is, benefits are given without the donor or the recipient feeling the recipient has an obligation to repay. This does not rule out the possibility that giving benefits increases the recipient's desire to behave communally toward the donor. It might and often does. Yet it might not. It simply means that the proximal motive for giving benefit is to improve the recipient's welfare and that neither the donor nor the recipient (if both were following a communal norm) feel that benefits come with a price tag, implicit or explicit. The donor may *hope* that the recipient will be similarly responsive to his or her needs as they arise which is likely why Goffman stated that something given in a social exchange, "need only be returned if the relationship calls for it." Yet, *hope* for the recipient having a similar motivation seems more appropriate to us than saying that a benefit "need only be returned when the relationship calls for it" because one can't demand such responsiveness. Moreover, we say *may hope* because there are communal relationships in which abilities to be responsive to one another's welfare differ greatly and, especially in these relationships, donors may not hope for similar responsiveness from the partner. For instance, parents may gladly pay their child's college tuition, yet if one of them decides to return to college they may well *not hope* that the child will strive to pay their tuition even if they must stretch to pay that extra tuition themselves. Ability to provide support matters (a factor that often explains why some communal relationships are asymmetrical in the sense that parties do not feel equal amounts of responsibility for one another's welfare – an aspect of communal relationships that entails differences in communal strength which is discussed in more detail below). Even when ability does not vary, some communal relationships may be asymmetrical and that may be OK with both sides.

Communal relationships are often exemplified by relationships commonly referred to as friends, family members, romantic partners, and spouses. Yet there are plenty of exceptions to this rule. Early on, when we first said that family relationships often exemplify communal relationships, a colleague responded, "Not my mother! She kept

careful track of what relatives gave us for wedding presents and then made sure her gift to their child was exactly comparable." Our sense was that this mother did have a communal relationship with her son but considered her relationships with many other relatives to be exchange in nature.

There is, we believe, an evolutionary basis for the existence of communal relationships. Newborn infants would not survive without someone attending to their needs non-contingently. Kin have likely long supported one another on a non-contingent basis. The very nature of small hunter-gatherer societies was such that there was an unpredictability of who would find food, shelter, and other necessities of life and who would need it. This likely dictated communal sharing and consumption of benefits (cf. Clark, 1984a; Clark and Jordan, 2002; and see Chapter 18, "Twists of Fate" in Kelley et al., 2003). Developmentally, an infant's need for and (one-sided) understanding of a communal norm would seem to emerge prior to his or her need for and understanding of an exchange norm which, in current society, is needed to give and receive benefits from a widening group of nonfamily members and nonfriends (see also Pataki et al., 1994). We once observed a young child at a community pool request a bag of potato chips at a snack bar and happily walk away without paying and without apparent guilt. The attendant called after the child telling him he must pay. Here, we thought, was an example of a child who likely understood communal norms well given that he was given food, [p. 235 ↓] unconditionally, by family members, and had some learning left to do when it came to understanding an exchange norm. Historically, the need for exchange in addition to communal relationships likely expanded greatly with the advent of civilization and specialization in skills, which increased differentiation between the skills and goods individuals could supply to one another.

Not all Relationships are Communal or Exchange in Nature

Although we have long focused upon communal and exchange relationships, we did not (and do not) believe that all relationships must be communal or exchange in nature. For example, there are also exploitative relationships which fit neither our conceptual definition for communal nor our conceptual definition for exchange relationships.

Relationships that seem to be some sort of hybrid of a communal and an exchange relationship also exist. For instance, elementary school teachers are responsible for attending to many aspects of their young students' welfare, especially their need to learn. They do not expect direct repayment from those children nor do the children give much thought (if any) to repayments. But the obligations are circumscribed and the same teacher is paid for these services albeit by the school district or school in exchange for these services and would not provide the services without pay.

Early Studies Demonstrating the Validity of the Qualitative Distinction

In the 1970s, social psychologists studying interpersonal interactions typically were conducting studies of interactions between strangers. We recall Ellen Berscheid, commenting at the time, that in our zeal to maintain control and to conduct true experiments on social behavior we were all busy studying relationships between people who had never seen one another before and who never expected to see each other in the future. We knew little about interactions in the relationships that matter most to people – those with their friends, family members, and romantic partners. All that has changed rapidly over the last three and a half decades (see, for example, Clark and Lemay, 2010) but to understand our research it is important to understand the context *at the time* the original research was conducted.

Once we had drawn the qualitative distinction, our challenge was to devise a true experimental manipulation of whether our participants would desire a communal or exchange relationship (and would therefore behave differently depending upon experimental condition when interacting with a partner if our postulates were correct). To create desire for a communal relationship a target person had to be interpersonally attractive, the research participant had to be "in the market" for new communal relationships and the target had to be available and interested as well. In the absence of these factors, we thought an exchange relationship would be preferred. We settled on a strategy of bringing two previously unacquainted people together, one of whom was a confederate, the other a true participant. We recruited college students in their early years at a residential college figuring that such students having just been uprooted

from family and high school communities were typically “in the market” for new friends and, possibly, a new romantic partner. We selected a physically attractive, relaxed, engaging young woman whom we knew to be popular as our confederate figuring that she would be an appealing potential friend or romantic partner. Then we manipulated *her* availability and seeming interest in a new communal relationship. In our communal condition, we described her as new at the university and eager to form relationships (making her similar to and available to our participants). In our exchange condition, we described her as married and about to be picked up by her husband (making her dissimilar to and unavailable to our participants). [p. 236 ↓] This manipulation worked (see, for instance, Clark and Mills, 1979; Clark, 1986) and was used with minor modification in many of our initial experimental studies.

Do People in (or Desiring) Communal Relationships Shun Behaviors Which Suggest They or a Potential Partner May Be following an Exchange Norm?

Our inaugural studies focused on demonstrating that exchange behaviors would occur and be welcomed in our exchange conditions, but would be avoided and reacted to negatively (if they did occur) in our communal conditions. We *had* to start here because, at the time, equity theory was the dominant theory for explaining how people gave and received benefits in relationships. The overall extant assumption was that giving benefits created inequities, resulted in discomfort, and called for repayment (see Walster et al., 1978). We predicted that *only* if an exchange relationship was desired would people positively respond to being repaid for a favor they had given to a target; reactions to such a payment would be negative when a communal relationship was desired. This should occur, we reasoned, because acting in accord with an exchange norm would imply that a communal relationship was not desired.

Our participants were male. They encountered a female target with whom they were led to desire either an exchange or a communal relationship. Each participant and

confederate worked on a task. They were allowed to help one another. Tasks were assigned to enable the real participant to finish his task with materials to spare. The experimenter asked if he wished to give materials to the female. All did so and the confederate repaid him or did not with an extra credit point. Finally, under the guise of preparation for another task the male participants indicated their liking of the female confederate.

The results were clear. Repayment increased liking (relative to no repayment) in the exchange conditions; in contrast it reduced liking in the communal conditions. In a second study (Clark and Mills, 1979, Study 2) participants were female and we manipulated desire for a communal or an exchange relationship with a female target. This time participants received help or did not follow by a request to return a comparable benefit or no request. After receiving a benefit, receiving a request for a comparable benefit increased liking in our exchange conditions but decreased liking in our communal conditions. Finally, receiving a request for a benefit in the absence of having received one decreased liking among our exchange condition participants (where it presumably created a debt) but not among our communal condition participants.

These studies and other early studies showing, for instance, that people who are not repaid feel exploited when they desire an exchange relationship but not when they desire a communal one (Clark and Waddell, 1985) and that giving and receiving non-comparable benefits results in a relationship looking more like a friendship than does giving and receiving comparable benefits (Clark, 1981) were greeted with enough interest to be published. Yet they also generated skepticism.

Reviewers and audience members at conferences suggested an alternative explanation: perhaps people in our exchange conditions wanted an *immediate* balancing of accounts whereas participants in our communal conditions still kept track of benefits given and received, but were content to let accounts be balanced across time. This was possible. Yet, we did not "buy" this interpretation for many reasons. It left open many questions: *Why* should there be a difference in time course? Also, even if the projected time courses were different why would someone be liked less in our communal conditions for prompt repayment than for failing to repay? Moreover, in the initial studies reported in Clark and Mills (1979) it seemed unlikely that there

was simply more time to “balance the books” in the communal than in [p. 237 ↓] the exchange conditions for there was no guarantee that the relationship in the communal conditions would be ongoing. Furthermore, there are plenty of long-term business (exchange) relationships in which bills still must be paid promptly. Why couldn’t those debts wait as well? It also seemed impossible to us that people simply could not (even if they wished to) keep track of and balance all the myriad benefits that are given and received over time in intimate relationships. In addition, an exchange rule could not explain such things as, why parents would ever care for a severely handicapped child or why the spouses of many Alzheimers patients continue to care for their partners even after the partners can no longer reciprocate or even recognize them. Finally, and importantly, following an exchange rule simply does not afford the benefit of the unambiguous inferences of being caring (by donors) and of being cared for (by recipients) of benefits. We sensed that these feelings were terribly important in friendships, family relationships, and romantic relationships.

We, of course, could reason all we wished about why we were right. What was needed to convince our critics was evidence that people do not keep track of benefits in all relationships as would be necessary to maintain equity across time. Thus, three studies were conducted to demonstrate this (Clark, 1984a, b). In the first, desire for a communal or an exchange relationship was manipulated; in the second two, we contrasted the behavior of strangers working together with that of friends working together. In each, participants engaged in a task in which they and a partner were working together to find certain number sequences in a large matrix of numbers. For each sequence found the pair earned a monetary reward to be divided afterwards. The partner always started working in either red or black ink. Then it was the participant's turn. Pens of both colors were available. The participant chose one and our dependent measure was whether the participant chose to work with the same color pen (making it unclear, in the end, how many sequences each person found) or a different color pen (making it absolutely clear who had contributed what). The results of all three studies were clear. Strangers who expected to remain strangers kept track of benefits – the vast majority choosing to use a different color pen; people led to desire a communal relationship or those with an existing communal relationship did not – each time fewer than the 50 percent did so. (This effect also was conceptually replicated by Clark et al., 1989). Indeed, when trying to *form* a communal relationship people seemed to “bend over backwards” not to

choose a different color pen choosing one significantly less often than 50 percent of the time.

These results were important. If people who desire or have a communal relationship do not keep track of benefits in the moment they cannot be quietly keeping those inputs in mind across time to make sure the books even out in the end.³

Do People in Communal Relationships Follow a Communal Norm?

Having established that people led to expect a friendship or romantic relationship (or those having such a relationship) react negatively to and avoid exchange behaviors, we turned to demonstrating that people led to desire communal relationships behave in more communal ways than those led to desire exchange relationships. In one set of studies we showed that people led to desire communal relationships would keep track of a light that indicated that a partner in the next room was experiencing a need (even if they could do nothing about it) whereas those who desired an exchange relationship were less likely to do so (Clark et al., 1986, Study 1; see also Clark et al., 1989). We also demonstrated that attention to needs would take place in exchange relationships if the person who could attend to those needs knew that he or she would soon be in the same position as the [p. 238 ↓] partner and would want the partner to attend to his or her needs. In that case, repayments might well be needed. Indeed, when roles were to be reversed, attention to needs in our exchange condition rose to levels comparable with those observed in our communal conditions (whether or not role reversals were scheduled). In the communal conditions attention to needs did not vary with the other person's opportunity to repay (Clark et al., 1986, Study 2). Also fitting with findings of attending more to needs when communal relationships are desired or exist, we found and reported evidence of people responding more positively to a partner's expressions of emotion (which are signals of needs or lack thereof) when a communal rather than an exchange relationship is desired (Clark and Taraban, 1991; Yoo et al., 2011, Study 1). We also found that people express more emotion when a communal rather than an exchange relationship is desired (Clark and Finkel, 2005a; Clark et

al., 2001). Furthermore, the willingness to express negative emotions prospectively predicts college students forming more relationships over the course of a semester with previously unknown fellow students and with the establishment of more intimacy in the closest of those relationships (Graham et al., 2008). Meanwhile other researchers have reported on links between suppressing emotions and lower social support, less closeness to others, and lower social satisfaction in normatively close relationships (Srivastava et al., 2009), lower rapport in such relationships, poorer communication, and lower chances of relationship formation (Butler et al., 2003).

Of course, levels of helping and responsiveness to another's sad emotion ought to be higher when communal relationships are desired or exist than when exchange relationships are desired or exist and we demonstrated this in a study by Clark et al. (1987, Study 1). People were randomly assigned to desire a communal or exchange relationship by giving them time alone to view an attractive confederate's photo and questionnaire responses that led them to believe that the confederate was either married or single, new to campus, and thinking that doing the study might be a good way to meet people. Participants were then provided an opportunity to voluntarily help a fellow participant by performing a mundane task. Participants in the communal condition helped for longer time periods than did people in the exchange condition. In addition, knowing the other was sad significantly increased time spent helping in the communal but not in the exchange conditions. Other studies revealed that people's moods improve in reaction to having helped a potential partner when communal (but not exchange) relationships are desired (Williamson and Clark, 1989; 1992) and that people feel badly about not helping when communal (but not when exchange) relationships are desired (Williamson et al., 1996). Moreover, people desiring a communal relationship care how much time another spends in picking out a gift *for them*; people desiring an exchange relationship do not (Clark et al., 1998).

Adding a Quantitative Dimension to the Theory

Whereas we started with a qualitative distinction, early on we added a quantitative dimension to our theory – communal strength (Mills and Clark, 1982) and later we

developed a measure of the dimension of communal strength (Mills et al., 2004). Although we considered it necessary to begin by emphasizing the qualitative distinction in order to emphasize the very existence of communal relationships and their distinction from exchange relationships, we knew from the beginning that communal relationships vary in how much responsibility in terms of time, effort, or money a person takes on for another person (what we call communal strength). Clearly, for instance, most people feel more communal responsibility for their children than for their friends and will spend more time, effort, and money to benefit their child than their friends. (Consider, for instance, [p. 239 ↓] the fact that many people pay college tuition for their children and almost no one does so for a friend although both relationships are, qualitatively, communal in nature.). Our measure of communal strength (Mills et al., 2004) has been shown to tap a construct distinct from behavioral interdependence as measured by Berscheid et al.'s (1989) Relationship Close ness Inventory and distinct from a measure of liking for the partner. (Consider, for instance, that one may feel considerable communal responsibility for a cranky elderly relative whom one does not like very much or that one can feel considerable liking for an attractive, potential romantic partner whom one has just met but for whom one feels little communal responsibility.) The communal strength scale correlates with Rubin's (1970) Love Scale and has been shown to predict allocation of benefits to peers and diary reports of giving help to and receiving help from friends. (See also Monin et al., 2008, for an additional methodology for measuring communal strength and further evidence that communal strength is not the same construct as liking).

People have very low-strength communal relationships with many other people. For instance, many people if stopped by a stranger and asked the time or for simple directions will provide the requested service to meet almost any requestor's needs without expecting anything in return. They have higher-strength communal relationships with others such as friends and people often have extremely high-strength communal relationships with a single other person or a very select group of people such as children and a spouse or romantic partner. [Figure 38.1](#) depicts one person's set of hierarchically arranged communal relationships. The hierarchy of relationships appears along the x-axis, the degree of felt responsibility appears along the y-axis and the line running through the graph depicts the degree of responsibility felt for various people. The needs of a person high in a communal hierarchy take precedence over

equivalent (and sometimes nonequivalent) needs of a person lower in communal strength in the event of a conflict. For instance, a person might forego attending a friend's birthday party in order to be at her own child's birthday party and the friend is very likely to understand and to accept this. Of course, [Figure 38.1](#) depicts just one possible hierarchy. The ordering of people (and of the self) in these implicit hierarchies will vary considerably between persons and cultures.

Figure 38.1 Costs one hypothetical person would be willing to incur to benefit a variety of relationship partners

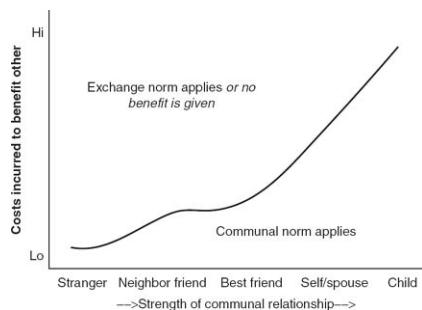
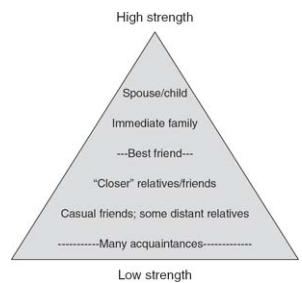


Figure 38.2 One person's hypothetical hierarchy of communal relationship partners



[p. 240 ↓] Reis et al. (2004) suggested that people's sets of communal relationships likely fall in a triangular shape as depicted in [Figure 38.2](#) with most people having many low-strength communal relationships, fewer medium-strength communal relationships, and very few very-high-strength communal relationships. There is, after all, a limit on the number of people for whom an individual can assume responsibility. It is also the case that once one has an established network of communal family members, friends

and a romantic partner who feel responsibility for one's own welfare, adding more adds little to one's own security.

Note that the “self” is also depicted in [Figures 38.1](#) and [38.2](#) because people do generally feel a communal responsibility to the self as well as to others. Most people place the “self” high in their hierarchy of communal relationships and consider their responsibility to care for themselves high and to take precedence over their responsibility to many other (but very often not all) other communal partners. For instance, a woman might think that taking care of herself takes clear precedence over taking care of the needs of a casual friend but not over taking care of her child.

Note also that [Figure 38.1](#) implicitly conveys that a person may have both a communal and an exchange relationship with the same relationship partner within different “cost” ranges. For instance, a person might give a friend advice, a ride, presents, lunch, and include that friend in social events all on a communal basis but sell the friend a car on an exchange basis. More commonly, we believe, benefits are given on a communal basis up to some “cost” threshold and above that they simply are not given or discussed.

Whereas the communal nature of a relationship has a quantitative dimension, exchange norms do not have an equivalent quantitative dimension (Mills and Clark, 1982). If a relationship is exchange in nature, it just is. Benefits given require comparable benefits in return. Repayments may be missed. After all these norms are ideals; violations will occur. For instance, given limited resources and many debts, some people may be forced to choose who to repay and who not to repay. They may do so based on liking, the importance of the exchange relationship to their well-being, the length of the relationship, and/or the demandingness of the person to whom the debt is owed. Yet unpaid debt remains, and with it comes guilt on the debtor's part and annoyance or anger on the grantor's part. There is no parallel in an exchange relationship to a friend's understanding that one could not attend to his need because one's child had a need at the same time. People may try to excuse unpaid debts in exchange relationships by appealing to having other debts but the person who is unpaid will just not understand.

The Distinction is Not One of Short- versus Long-Term Relationships

From the time we published our first paper on the distinction between exchange versus communal relationships, many people made two assumptions about its basis, neither of which we shared. Daniel Batson raised both in a paper challenging our distinction (Batson, 1993). He actually did us a favor in the sense of pushing us to address in print two issues that others had raised as well (see Clark and [p. 241 ↓] Mills, 1993 for our response). The first assumption he made was that exchange relationships are short-term relationships with each benefit given being quickly repaid, whereas communal relationships are long-term relationships with benefits balancing out across time. As we have already noted in talking about record keeping, this is not the basis for our distinction. Communal and exchange relationships can be either short- or long-term. Some (generally weak) communal relationships are one-time occurrences as in our earlier example of a person telling a stranger the time without expecting anything in return. Other times (as in some emergency situations) one can provide another with larger communal benefits on a short-term basis as once occurred to one of us when the parent of an incoming college student had no money for housing upon bringing her daughter to school and stayed (along with her pets and daughter) in our home for awhile. Other communal relationships, such as many marriages, parent-child relationships, and friendships do last a long time. Responsiveness to needs and desires occurs as those needs and desires arise across time. In symmetrical communal relationships, benefiting may well be roughly equal across time. Yet there are no guarantees and there are also many cases of asymmetrical long-term communal relationships in which benefits may never even out. Perhaps the most common examples are the relationship between parents and children during the younger individuals' childhoods and early adulthood. Of course, if the parents need help as they age it is often given but both parties are unlikely to be distressed (and in fact very likely to be happy) if that need never arises.

The Distinction is Not One of Selfish versus Unselfish Relationships

A second assumption Batson (1993) and others have made is that communal relationships are unselfish in nature and exchange relationships selfish. We do not make that assumption. Once a communal norm is adopted, benefits are given on a non-contingent basis; once an exchange norm is adopted, benefits are given on a contingent basis but *either* selfish or unselfish motives can drive a person to adopt each norm.

Consider a communal norm, for instance. There exist many possible “selfish” reasons for adopting such a norm. One may have just moved to a new community, wish to form new friendships, and act on a communal basis toward potential friends to start a friendship. One may care for a disagreeable, elderly relative on a communal basis because one would feel guilty if one did not or because one fears criticism by others. There also exist unselfish reasons for following a communal norm such as feelings of empathy for one's partner. So too may the drive to communally care for one's offspring compel one to adherence to a communal norm in a manner that seems unselfish to us (though some might consider it selfish in the sense of promoting the survival of one's genes across generations).

The adoption of an exchange norm also may be driven by relatively selfish or relatively unselfish motives. It is likely that most times when a person adopts an exchange norm the motive is selfish. For example, when a person goes to the store to buy a loaf of bread it is because he wants that bread and almost never because he wants to benefit the grocer. When a person forms a car pool it is likely to save time and money for him or herself. Yet people may also adopt an exchange norm for unselfish reasons. It might, for instance, be possible to exploit an employee given a dearth of jobs and to pay the person less than his or her work is worth. If that were done the relationship would not adhere to an exchange norm but would be best characterized as exploitative in nature. In such a situation, unselfish motives (to be moral, to be fair) might drive the decision to follow an exchange norm by paying the person a fair wage. Of course, both selfish and unselfish motives might drive adoption of an exchange (or a communal) relationship.
[p. 242 ↓] For instance, currently people advocate buying “fair trade” coffee on an

exchange basis seemingly *both* because they want coffee (a selfish motive) and because, although equally good coffee might be purchased at a lower price, they do not wish to exploit coffee workers and instead want to offer them a “fair” trade in exchange terms (a relatively unselfish motive although we know that some might say the buyer gets to feel good about himself and that makes it selfish).

In discussing whether communal relationships are selfish or not, it is also important to point out our assumption that people place themselves in their own hierarchies of communal relationships and that, typically, they place themselves high in those hierarchies. This means that most people do consider taking care of their own needs to take precedence over taking care of most other people's needs even when it may be said that they have communal relationships with those others. It is easy to illustrate this point. People send themselves, not their neighbors or friends, on vacation, but they still can act communally toward those friends.

They may also place certain others at a rank equal to the self in their hierarchies (e.g., a spouse may merit this rank) and some at higher ranks in their hierarchies (e.g., a young, completely dependent child is often “given” such a rank) and make sacrifices for such people and or forgive such people for major wrongdoings. The overall point to be made here, though, is simply that selfishness versus unselfishness is not the defining characteristic of communal relationships.

Communal Relationships Can Be Symmetrical or Asymmetrical

Most communal relationships are symmetrical in the sense that each person assumes about the same level of responsibility for the partner's welfare, as does the partner for him or her. Friendships, romantic relationships, and marriages are examples of relationships that are often both communal in nature and symmetrical in felt responsibility. Yet communal relationships can be asymmetrical as well. A mother generally assumes much greater responsibility for her young child's welfare than vice versa and this pattern often continues right into the child's adulthood. Indeed, one of us recently asked a large class of college students about whether their relationships

with their friends and their relationships with their mothers were characterized by feeling “about equal responsibility for one another’s welfare,” the other feeling “more responsibility for me than I do for her/him,” or the self feeling “more responsibility for her/him than s/he does for me.” For friends, 84 percent of the students reported feeling equal responsibility, 6 percent said their friend felt more responsibility for them than they did for the friend, and 10 percent said they felt more responsibility for their friend than vice versa. Reports for their mother showed a very different pattern. Only 15 percent said they and their mother felt equal responsibility for one another, 85 percent reported their mother felt more responsibility for them than vice versa, and no student reported that he or she felt more responsibility for his or her mother than she did more the student. Clearly these particular Western, largely affluent, college students tended to have symmetrical communal relationships with their friends and asymmetrical ones with their mothers.

Individual Differences in Communal and Exchange Orientation

We believe most variability in communal responsiveness lies between relationships rather than between individuals (Clark and Lemay, 2010). That is, we assume that almost all people have relationships in which they strive to follow a communal norm as well as other relationships in which they make little or no effort to be responsive to partners. That said, differences in people's tendencies [p. 243 ↓] to follow communal or exchange norms also exist.

Box 38.1 Scales to Measure Individual Differences in Communal and in Exchange Orientation

Communal Orientation Scale

- 1. It bothers me when other people neglect my needs.
- 2. When making a decision, I take other people's needs and feelings into account.
- 3. I'm not especially sensitive to other people's feelings.*
- 4. I don't consider myself to be a particularly helpful person.*
- 5. I believe people should go out of their way to be helpful.
- 6. I don't especially enjoy giving others aid.*
- 7. I expect people I know to be responsive to my needs and feelings.
- 8. I often go out of my way to help another person.
- 9. I believe it's best not to get involved taking care of other people's personal needs.*
- 10. I'm not the sort of person who often comes to the aid of others.*
- 11. When I have a need, I turn to others I know for help.
- 12. When people get emotionally upset, I tend to avoid them.*
- 13. People should keep their troubles to themselves.*
- 14. When I have a need that others ignore, I'm hurt.

Items from the Exchange Orientation Scale

- 1. When I give something to another person, I generally expect something in return.
- 2. When someone buys me a gift, I try to buy that person as comparable a gift as possible.
- 3. I don't think people should feel obligated to repay others for favors.*
- 4. I wouldn't feel exploited if someone failed to repay me for a favor.*
- 5. I don't bother to keep track of benefits I have given others.*
- 6. When people receive benefits from others, they ought to repay those others right away.
- 7. It's best to make sure things are always kept 'even' between two people in a relationship.
- 8. I usually give gifts only to people who have given me gifts in the past.
- 9. When someone I know helps me out on a project, I don't feel I have to pay them back.*

* Note: These are two independent scales. Respondents rate each item for each scale on a five-point scale from "extremely uncharacteristic" of them (1) to "extremely

characteristic" of them (5). Scores for items followed by an asterisk are reversed prior to calculating a sum indicating the respondent's communal or the respondent's exchange score.

To measure such individual differences, scales of both communal orientation and of exchange orientation have been developed (Clark et al., 1987; Mills and Clark, 1994). These two separate and orthogonal scales appear in [Box 38.1](#). The communal scale has been utilized more extensively than the exchange scale (but note that other researchers, for instance Murstein and his colleagues [1977] and Sprecher [1992], have developed their own methods for measuring exchange orientation and exchange orientation generally has been studied a fair amount).

High scores on communal orientation (using the scale shown in [Box 38.1](#)) have been shown to predict: helping a fellow student in a nonemergency situation (Clark et al., 1987); agreement that support has taken place among friends (Coriell and Cohen, 1995); willingness to express emotion to relationship partners especially when the relationship context calls for so doing (Clark and Finkel, 2005a, 2005b); allocating rewards equally when negotiating a friend, which makes good sense communally assuming equal needs (Thompson and Deharppot, 1998); people giving partners more credit for joint successful performances on a task and blaming them less for failure while attributions to the self remain unaffected by communal orientation (McCall, 1995); greater satisfaction in elderly persons' best friendships (Jones and [p. 244 ↓] Vaughan, 1990); and responding to cues of power with greater social responsibility (Chen et al., 2001). Low scores predict: burnout in nurses (Van Yperen et al., 1992) and leaders of self-help groups (Medvene et al., 1997); depression among caregivers of Alzheimer's patients (Williamson and Schulz, 1990); male physical abuse of female partners as well as males' associating with peers who endorse violence against female partners (Williamson and Silverman, 2001); and links between feeling under-benefited and feeling resentful (Thompson et al., 1995). *Matches* on communal orientation have been shown to be linked to better ability to capitalize on mutual opportunities in negotiations (Thompson and Deharppot, 1998). In other words, individual differences in communal orientation can be measured and predict behaviors that one would expect on the basis of following (or failing to follow) a communal norm.

Box 38.2 The Ten-Item Communal Strength Measure

- 1. How far would you be willing to go to visit_____?
- 2. How happy do you feel when doing something that helps_____?
- 3. How large a benefit would you be likely to give_____?
- 4. How large a cost would you incur to meet a need of_____?
- 5. How readily can you put the needs of _____ out of your thoughts?*
- 6. How high a priority for you is meeting the needs of_____?
- 7. How reluctant would you be to sacrifice for_____?*
- 8. How much would you be willing to give up to benefit_____?
- 9. How far would you go out of your way to do something for_____?
- 10. How easily could you accept not helping_____?

* Note: The items with asterisks are reversed prior to summing scores. The instructions for this scale are as follows: "As you answer each question, fill in the person's initials in the blank. Circle one answer for each question on the scale from 0 'not at all' to 10 'extremely' before going on to the next question. Your answers will remain confidential."

High exchange orientation (using a variety of measures) has been found to predict one's marital satisfaction being tied to considerations of equity and, overall, to lower marital satisfaction, whereas the marital satisfaction of people low in exchange orientation has been shown to be both higher and unrelated to considerations of equity (Buunk and Van Yperen, 1991; see also Murstein et al., 1977) as well as to expectations of becoming distressed over inequities in a relationship (Sprecher, 1992). High exchange orientation also has been linked to lower compatibility and friendship ratings among roommates and higher anxiety among women in those roommate pairs (Murstein and Azar, 1986).

Many situations clearly call for following one norm or the other and most people, no matter what their overall orientations, adhere to the norm that matches the situation. Yet individual differences in orientation likely do come into play in two types of situations. First, they are likely to come into play in situations lacking strong situational cues regarding how to behave. For instance, Clark et al. (1987, Study 2) found that

communal orientation scores predicted how much help a person gave a young research assistant when no manipulation of relationship type had taken place. Second, these individual differences likely influence how easily and with how much equanimity people are able to follow the norm appropriate to a particular situation. That is, we suspect that exchangeoriented individuals likely must exert more effort and more self-consciously follow communal norms in marriage than others. They may also violate a communal norm/appeal to an exchange norm more often than others. In addition, we suspect that communally oriented individuals may have a tougher time [p. 245 ↓] sticking to exchange norms when the situation calls for using that norm but evidence of a person being needy exists. For instance, they might find it especially difficult to terminate an employee who needs the job but has not been living up to the (exchange) terms of employment.

Recent work also has begun examining how the individual difference dimensions of attachment-related avoidance and anxiety relate to adherence to communal and to exchange norms. Evidence suggests that higher levels of attachment-related avoidance are linked to: some reluctance to enter situations in which one may (or may not) receive evidence of another's communal interest; slightly greater tendencies to behave in accord with an exchange norm when in situations that are normatively communal in nature; and discomfort in situations in which adherence to a communal norm is occurring or seems to be called for (cf. Bartz and Lydon, 2006, 2008; Beck and Clark, 2009; Clark et al., 2010). Higher levels of attachment related anxiety do not seem to be linked to reluctance to enter communal situations but do seem to be related to ambivalence and discomfort in such situations and more reactivity to behavior suggesting adherence or lack thereof to both communal and exchange norms (Bartz and Lydon, 2006, 2008; Clark et al., 2010).

Four Additional Theoretical Points

Our work on communal relationships has led us to postulate a few criteria for the existence of "high quality" close relationships in addition to being responsive to partner desires and needs and seeking such responsiveness from partners. First, because most peer communal relationships are symmetrical, couple members who implicitly agree on the appropriate strength of their communal relationship and its desired trajectory ([Box](#)

38.2) (both in terms of the slope and speed of strengthening the nature of the communal relationship or the sense that the relationship should not be strengthened) ought to feel more satisfied and comfortable with their relationships than those who do not agree. If one person desires a stronger communal relationship than the other, first person may feel neglected and latter partner may feel smothered.

Second, in connection with people's (usually implicit) hierarchy of communal relationships, we suggest that a couple's or friends' complementarity in communal hierarchies will influence the quality of their relationships with one another. For example, spouses who are in agreement with one another that their newborn's welfare takes precedence over both of their own needs, their obligations to one another come next, and their obligations to their respective families of origin rank third, ought to experience less conflict in their relationship than a couple including a wife who puts her infant first, her parents second and her spouse third while the spouse puts her first (and expects her to do the same for him), his child second, and his own family of origin third (and, expects her to do the same with regard to their child and her family of origin).

Third, we believe that placement of the self within one's hierarchy of communal relationships has important implications for the nature of one's strongest communal relationships and, indeed, to one's ability to have very strong communal relationships. In particular, placing the self high in one's hierarchy but having another person (e.g., a spouse, a child) placed higher in one's hierarchy or at least "tied" with the self may be a requirement for "pulling off" very strong communal relationships. Here is why.

When the self is placed alone at the top of the hierarchy, especially when the self is placed well above everyone else, attending to the self's own needs will always take precedence over attending to others' needs. Compromises and sacrifices will not be made. Forgiveness for transgressions will not take place. Partners will not be able to relax self-defenses knowing that there is someone else to care for them as much or more than they care for themselves.

Finally, we do not suggest that adopting a communal norm is always the best strategy [p. 246 ↓] in relationships nor always healthy. Adopting a communal norm when the partner prefers an exchange relationship can be awkward and distressing to both parties. Similarly, acting in accord with a strong communal norm when one's partner

desires a weaker communal relationship can produce problems as can mismatches in desires for asymmetrical versus symmetrical communal relationships.

Do People Really Believe in and Follow Communal Norms in Their Intimate Relationships?

Do people really believe a communal rule is the “right” rule for their friendships, romantic relationships, and family relationship? Do they actually follow this rule in ongoing relationships? After all, one might think, most of the early research was done utilizing relationships between people who were meeting for the first time. Perhaps such people do follow a communal rule to win one another’s affections but then drop the rule after commitments have been made. Recently we have been studying ongoing marriages and the results for this research suggest the answer to the first question posed above is yes. People do believe a communal norm is ideal for their marriages and that an exchange norm is decidedly not ideal, and in at least two samples of marriages the vast majority of people report that they and their spouse strive to follow a communal norm. The answer to the second question also appears to be yes (with some caveats). Specifically, both Grote and Clark (1998) and Clark et al. (2010) have found that individuals will rate a communal norm as ideal and an exchange norm as decidedly not ideal for their marriages. Clark et al. (2010) have also shown that, for at least the first two years of marriage, spouses report striving to adhere to such a norm, and that their partners strive to adhere to this norm as well. The caveats are straightforward. Although members of both samples overwhelmingly reported that both they and their spouse strive to follow communal norms, research also suggests that, especially when they are distressed (Grote and Clark, 2001) or have chronic relationship insecurities (Clark et al., 2010), they may “fall down” on the job and calculate fairness according to an exchange rule.

Of course, it is also the case that the clinical and counseling literature on relationships provides overwhelming evidence that some relationships people normatively desire to be communal in nature (e.g., marriages) may come to be characterized by the antithesis

of communal caring. Members may verbally and physically abuse one another, berate and criticize one another, and show contempt for one another. Although we have consistently said that family relationships, romantic relationships and marriages often exemplify communal relationships, at times, we hasten to add, they definitely do not. Adherence to a communal norm, we believe, characterizes well-functioning, healthy, marriages, friendships, and family relationships. Therapists will certainly encounter people in marriages and friendships and family relationships who do not follow communal norms. That ought not to be taken as evidence of a communal norm not applying to intimate relationships or of there being no differences in rules that govern the giving and acceptance of benefits in different relationships. Appeals to an exchange norm may be made in relationships that society calls upon to be communal in nature. When they are made, we suggest, they are signs of trouble. In contrast, when exchange norms are appealed to and followed in, say, a business relationship, they suggest the relationship is a healthy one. A person would not seek counseling because his business partner is keeping track of just who contributes what to the business and just who derives what benefits from the business. If his wife or mother or best friend did the same, it would not be surprising for counseling to be sought.

[p. 247 ↓]

Conclusion

When we set forth our distinction between communal and exchange relationships, we were forging new ground in social psychology. That has all changed dramatically as we pass the thirtieth anniversary of our first communal/exchange paper (Clark and Mills, 1979). Now work on relationships that are, normatively, communal in nature is thriving. Much knowledge has been gained about intra- and interpersonal processes characteristic of well and poorly functioning relationships that society calls for to be communal in nature. We have not attempted a review of the extensive work regarding factors that promote the healthy communal functioning of close relationships and the factors that interfere with such relationships (but see Clark and Lemay, 2010, for a review of such work, much of it done by others; and Clark et al., 2010; Grote and Clark, 2001; Grote et al., 2002, 2004; Lemay and Clark, 2008; Lemay et al., 2007 for theoretical ideas and results from our own laboratory suggesting factors that contribute

to and detract from the communal health of a communal relationship.). Rather we have tried to convey a good sense of our original qualitative distinction between communal and exchange relationships, the quantitative concept of the communal strength of relationships, some implications of our theory for relationship functioning and of the empirical research that we, ourselves, have done to test our theoretical ideas.

We conclude by returning to the questions we posed to our readers at the beginning of this chapter. First, why, when purchasing a gift for a friend, do we expect price tags to be on items yet after the purchase make sure they been removed? The answer is that relationships with friends are communal in nature and relationships with storeowners are exchange in nature. Second, why did the friend who rented a vacation house and arrived to find no hot water find it maddening that her real estate agent tried to elicit her sympathy by explaining that the owners were experiencing severe personal problems? It is because a rental arrangement is exchange in nature and the excuse called upon her to feel communal understanding for the unknown owners with whom she had no relationship at all.

Acknowledgments

Much of the work reported here was supported by NSF grant BNS 9983417. Opinions, conclusions and recommendations expressed in this paper, however, are those of the authors and do not necessarily reflect the views of the National Science Foundation. I thank Oriana Aragon and Elizabeth Clark-Polner for detailed comments on this chapter.

Notes

1 After having been invited to prepare this chapter but prior to its completion, Judson R. Mills died. He was my mentor, the inspiration for the original communal/exchange distinction, central to the theoretical ideas expressed here and a co-author on much of the empirical work. In this chapter I have tried to stay true to his thinking. Yet had he lived he would forced me to be conceptually clearer and more precise, added new conceptual ideas, and most certainly debated the new ideas expressed here.

2 In excluding rewards from our theory, we immediately were addressing a more narrow set of issues than equity theorists had dealt with for many equity theorists had included rewards in their calculation of equity (cf. Walster et al., 1978).

3 In saying this it may be helpful to keep in mind that this does not mean that we believe people never violate a communal norm and keep track nor that when people feel their needs have been neglected that they sometimes retrospectively try to calculate "fairness." They do (Grote and Clark, 2001) – often in a very biased manner! It does mean that doing so is a violation of a norm typically followed in communal relationships.

References

Adams, J.S. Inequity in social exchange . Advances in Experimental Social Psychology , 62, 335–343. (1965)

Bartz, J.A. and Lydon, J.E. Navigating the interdependence dilemma: Attachment goals and the use of communal norms with potential close others . Journal of Personality and Social Psychology , 9, 77–96. (2006) <http://dx.doi.org/10.1037/0022-3514.91.1.77>

Bartz, J.A. and Lydon, J.E. Relationship specific attachment, risk regulation, and communal norm adherence in close relationships . Journal of Experimental Social Psychology , 4, 655–663. (2008) <http://dx.doi.org/10.1016/j.jesp.2007.04.003>

Batson, C.D. Communal and exchange relationships: What is the difference? Personality and Social Psychology Bulletin , 19, 677–683. (1993) <http://dx.doi.org/10.1177/0146167293196002>

Beck, L.A. and Clark, M.S. Choosing to enter or avoid diagnostic social situations . Psychological Science , 20, 1175–1181. (2009) <http://dx.doi.org/10.1111/j.1467-9280.2009.02420.x>

Berscheid, E., Snyder, M. and Omoto, A. The relationship closeness inventory: Assessing the closeness of interpersonal relationships . Journal of Personality and Social Psychology , 57, 792–807. (1989) <http://dx.doi.org/10.1037/0022-3514.57.5.792>

Butler, E.A., Egloff, B., Wilhelm, F.H., Smith, N.C., Erickson, E.A. and Gross, J.J. The social consequences of expressive suppression . *Emotion* , 3, 48–67. (2003) <http://dx.doi.org/10.1037/1528-3542.3.1.48>

Buunk, B.P. and Van Yperen, N.W. Referential comparisons, relational comparisons, and exchange orientation: Their relation to marital satisfaction . *Personality and Social Psychology Bulletin* , 17, 709–717. (1991) <http://dx.doi.org/10.1177/0146167291176015>

Chen, S., Lee-Chai, A.Y. and Bargh, J.A. Relationship orientation as a moderator of the effects of social power . *Journal of Personality and Social Psychology* , 80, 173–187. (2001) <http://dx.doi.org/10.1037/0022-3514.80.2.173>

Clark, M.S. Non-comparability of benefits given and received: A cue to the existence of friendship . *Social Psychology Quarterly* , 44, 375–381. (1981) <http://dx.doi.org/10.2307/3033907>

Clark, M.S. (1984a) A distinction between two types of relationships and its implications for development . In J.C. Masters, ed. and K. Yarkin-Levin (eds), *Boundary Areas in Social Psychology* , pp. 241–270. New York: Academic Press.

Clark, M.S. Record keeping in two types of relationships . *Journal of Personality and Social Psychology* , 47, 549–557. (1984b) <http://dx.doi.org/10.1037/0022-3514.47.3.549>

Clark, M.S. (1985) Implications of relationship type for understanding compatibility . In W. Ickes (ed.), *Compatible and Incompatible Relationships* , pp. 199–140. New York: Springer-Verlag.

Clark, M.S. Evidence for the effectiveness of manipulations of desire for communal versus exchange relationships . *Personality and Social Psychology Bulletin* , 12, 414–425. (1986) <http://dx.doi.org/10.1177/0146167286124004>

Clark, M.S., Dubash, P. and Mills, J. Interest in another's consideration of one's needs . *Journal of Experimental Social Psychology* , 34, 246–264. (1998) <http://dx.doi.org/10.1006/jesp.1998.1352>

Clark, M.S. and Finkel, E.J. Willingness to express emotion: The impact of relationship type, communal orientation and their interaction . *Personal Relationships* , 12, 169–180. (2005a) <http://dx.doi.org/10.1111/j.1350-4126.2005.00109.x>

Clark, M.S. and Finkel, E.J. (2005b) Does expressing emotion promote well-being? It depends on relationship context . In L.A. Tiedens, ed. and C.W. Leach (eds), *The Social Life of Emotions* , pp. 105–126. New York: Cambridge University Press.

Clark, M.S., Fitness, J. and Brissette, I. (2001) Understanding people's perceptions of relationships is crucial to understanding their emotional lives . In G. Fletcher, ed. , and M.S. Clark (eds), *Blackwell Handbook of Social Psychology*, Vol. 3: *Interpersonal Processes* , pp. 253–278. Oxford, UK: Blackwell.

Clark, M.S. and Jordan, S. (2002) Adherence to communal norms: What it means, when it occurs and some thoughts on how it develops . In B. Laursen, ed. and W. Graziano (eds), *Justice: New Directions for Child and Adolescent Development* , pp. 3–25. New York: Jossey-Bass. <http://dx.doi.org/10.1002/cd.34>

Clark, M.S. and Lemay, E. (2010) Close Relationships . In S.T. Fiske, ed. , D.T. Gilbert, ed. and G. Lindsey (eds), *Handbook of Social Psychology* , Vol. 2, 5th Edition, pp. 898–940. Hoboken, NJ: John Wiley & Sons, Inc.

Clark, M.S., Lemay, E.P., Graham, S.M., Pataki, S. and Finkel, E.J. Ways of giving and receiving benefits in marriage: What's ideal? What happens? Attachment style matters . *Psychological Science* , 21, 944–951. (2010) <http://dx.doi.org/10.1177/0956797610373882>

Clark, M.S. and Mills, J. Interpersonal attraction in exchange and communal relationships . *Journal of Personality and Social Psychology* , 37, 12–24. (1979) <http://dx.doi.org/10.1037/0022-3514.37.1.12>

Clark, M.S. and Mills, J. (2001) Behaving in such a way as to maintain and enhance relationship satisfaction . In J.H. Harvey, ed. and A.E. Wenzel (eds), *Relationship Maintenance and Enhancement* , pp. 13–26. Mahwah, NJ: Erlbaum.

Clark, M.S., Mills, J. and Corcoran, D. Keeping track of needs and inputs of friends and strangers . *Personality and Social Psychology Bulletin* , 15, 533–542. (1989) <http://dx.doi.org/10.1177/0146167289154007>

Clark, M.S., Mills, J. and Powell, M. Keeping track of needs in two types of relationships . *Journal of Personality and Social Psychology* , 51, 333–338. (1986) <http://dx.doi.org/10.1037/0022-3514.51.2.333>

Clark, M.S., Ouellette, R., Powell, M. and Milberg, S. Recipient's mood, relationship type, and helping . *Journal of Personality and Social Psychology* , 53, 94–103. (1987) <http://dx.doi.org/10.1037/0022-3514.53.1.94>

Clark, M.S. and Taraban, C.B. Reactions to and willingness to express emotion in two types of relationships . *Journal of Experimental Social Psychology* , 27, 324–336. (1991) <http://dx.doi.org/10.1016/0022-1031%2891%2990029-6>

Clark, M.S. and Waddell, B. Perceptions of exploitation in communal and exchange relationships . *Journal of Social and Personal Relationships* , 2, 403–418. (1985) <http://dx.doi.org/10.1177/0265407585024002>

Coriell, M. and Cohen, S. Concordance in the face of a stressful event: When do members of a dyad agree that one person supported the other? *Journal of Personality and Social Psychology* , 69, 289–299. (1995) <http://dx.doi.org/10.1037/0022-3514.69.2.289>

Goffman, I. (1959) *The Presentation of Self in Everyday Life* . Garden City: Doubleday and Company.

Graham, S.M., Huang, J., Clark, M.S. and Helgeson, V. The positives of negative emotion: Willingness to express negative emotions promotes relationships . *Personality and Social Psychology Bulletin* , 34, 394–406. (2008) <http://dx.doi.org/10.1177/0146167207311281>

Grote, N.K. and Clark, M.S. Distributive justice norms and family work: What is perceived as ideal, what is applied, and what predicts perceived fairness? *Social Justice Research* , 11, 243–269. (1998) <http://dx.doi.org/10.1023/A:1023234732556>

Grote, N. and Clark, M.S. Does conflict drive perceptions of unfairness or do perceptions of unfairness drive conflict? *Journal of Personality and Social Psychology*, 80, 281–293 362 (erratum). (2001) <http://dx.doi.org/10.1037/0022-3514.80.2.281>

Grote, N.K., Clark, M.S. and Moore, A. Perceptions of injustice in family work: The role of psychological distress . *Journal of Family Psychology* , 18, 480–492. (2004) <http://dx.doi.org/10.1037/0893-3200.18.3.480>

Grote, N.K., Naylor, C. and Clark, M.S. Perceiving the division of family work to be unfair: Do social comparisons, enjoyment, and competence matter? *Journal of Family Psychology* , 16, 510–522. (2002) <http://dx.doi.org/10.1037/0893-3200.16.4.510>

Jones, D.C. and Vaughan, K. Close friendships among senior adults . *Psychology and Aging* , 3, 451–457. (1990) <http://dx.doi.org/10.1037/0882-7974.5.3.451>

Kelley, H.H., Hollmes, J.G., Kerr, N.L., Reis, H.T., Rusbult, C E. and Van Lange, P.A.M. (2003) *An Atlas of Interpersonal Situations* . Cambridge: Cambridge University Press.

Lemay, E.P. and Clark, M.S. How the head liberates the heart: Projection of communal responsiveness guides relationship promotion . *Journal of Personality and Social Psychology* , 94, 647–671. (2008) <http://dx.doi.org/10.1037/0022-3514.94.4.647>

Lemay, E.P., Clark, M.S. and Feeney, B. Projection of responsiveness to needs and the construction of satisfying communal relationships . *Journal of Personality and Social Psychology* , 92, 834–853. (2007) <http://dx.doi.org/10.1037/0022-3514.92.5.834>

McCall, M. Orientation, outcome and other-serving attributions . *Basic and Applied Social Psychology* , 17, 49–64. (1995)

Medvene, L.I., Volk, F.A. and Meissen, G.J. Communal orientation and burnout among self-help group leaders . *Journal of Applied Social Psychology* , 27, 262–278. (1997) <http://dx.doi.org/10.1111/j.1559-1816.1997.tb00632.x>

Messick, D. and Cook, K. (1983) *Equity Theory: Psychological and Sociological Perspectives* . New York: Prager.

Mills, J. and Clark, M.S. (1982) Exchange and communal relationships . In L. Wheeler (ed.), *Review of Personality and Social Psychology* , pp. 121–144. Beverly Hills, CA: Sage.

Mills, J. and Clark, M.S. Communications that should lead to perceived exploitation in communal and exchange relationships . *Journal of Social and Clinical Psychology* , 4, 225–234. (1986) <http://dx.doi.org/10.1521/jscp.1986.4.2.225>

Mills, J. and Clark, M.S. (1994) Communal and exchange relationships: New research and old controversies . In R. Gilmour, ed. and R. Erber (eds), *Theoretical Approaches to Personal Relationships* , pp. 29–42. Hillsdale, NJ: Erlbaum.

Mills, J. and Clark, M.S. (2001) Viewing close romantic relationships as communal relationships: Implications for maintenance and enhancement . In J. Harvey, ed. and A. Wenzel (eds), *Close Romantic Relationships: Maintenance and Enhancement* , pp. 13–25. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Mills, J., Clark, M.S., Ford, T. and Johnson, M. Measurement of communal strength . *Personal Relationships* , 11, 213–230. (2004) <http://dx.doi.org/10.1111/j.1475-6811.2004.00079.x>

Monin, J.K., Clark, M.S. and Lemay, E. Communal responsiveness in relationships with female versus male family members . *Sex Roles* , 59, 176–188. (2008) <http://dx.doi.org/10.1007/s11199-008-9420-8>

Murstein, B.I. and Azar, J.A. The relationship of exchange-orientation to Friendship intensity, roommate compatibility, anxiety, and friendship . *Small Group Research* , 17, 3–17. (1986) <http://dx.doi.org/10.1177/104649648601700101>

Murstein, B.I., Cerreto, M. and MacDonald, M.G. A theory and investigation of the effect of exchange-orientation on marriage and friendship . *Journal of Marriage and the Family* , 39, 543–548. (1977) <http://dx.doi.org/10.2307/350908>

Murstein, B.I., Wadlin, R. and Bond, C.F. The revised exchange-orientation scale . *Small Group Behavior* , 18, 212–223. (1987) <http://dx.doi.org/10.1177/104649648701800205>

Pataki, S., Shapiro, C. and Clark, M.S. Children's acquisition of appropriate norms for friendships and acquaintances . *Journal of Social and Personal Relationships* , 11, 427–442. (1994) <http://dx.doi.org/10.1177/0265407594113008>

Reis, H., Clark, M.S. and Holmes, J. (2004) Perceived partner responsiveness as an organizing construct in the study of intimacy and closeness . In D. Mashek, ed. and A. Aron (eds), *The Handbook of Closeness and Intimacy* . Mahwah, NJ: Lawrence Erlbaum Associates.

Rubin, Z. Measurement of romantic love . *Journal of Personality and Social Psychology* , 16, 265–273. (1970) <http://dx.doi.org/10.1037/h0029841>

Sprecher, S. How men and women expect to feel and behavior in response to Inequity in close relationships . *Social Psychology Quarterly* , 55, 57–69. (1992) <http://dx.doi.org/10.2307/2786686>

Srivastava, S., Tamir, M., McGonigal, K.M., John, O.P. and Gross, J.J. The social costs of emotional suppression: A prospective study of the transition to college . *Journal of Personality and Social Psychology* , 96, 883–897. (2009) <http://dx.doi.org/10.1037/a0014755>

Thibaut, J.W. and Kelley, H.H. (1959) *The Social Psychology of Groups* . New York: Wiley.

Thompson, L. and Deharport, T. Relationships, goal incompatibility, and communal orientation in negotiations . *Basic and Applied Social Psychology* , 20, 33–44. (1998) http://dx.doi.org/10.1207/s15324834basp2001_4

Thompson, S.C., Medvene, L.J. and Freedman, D. Caregiving in the close relationships of cardiac patients: Exchange, power, and attributional perspectives on caregiver resentment . *Personal Relationships* , 2, 125–142. (1995) <http://dx.doi.org/10.1111/j.1475-6811.1995.tb00081.x>

VanYperen, N., Buunk, B.P. and Schaufeli, W.B. Imbalance, communal orientation, and the burnout syndrome among nurses . *Journal of Applied Social Psychology* , 22, 173–189. (1992) <http://dx.doi.org/10.1111/j.1559-1816.1992.tb01534.x>

Walster, E., Walster, G.W. and Berscheid, E. (1978) Equity: Theory and Research . New York: Allyn & Bacon.

Williamson, G.M. and Clark, M.S. Providing help and desired relationship type as determinants of changes in moods and self-evaluations . Journal of Personality and Social Psychology , 56, 722–734. (1989) <http://dx.doi.org/10.1037/0022-3514.56.5.722>

Williamson, G.M. and Clark, M.S. Impact of desired relationship type on affective reactions to choosing and being required to help . Personality and Social Psychology Bulletin , 18, 10–18. (1992) <http://dx.doi.org/10.1177/0146167292181002>

Williamson, G.M., Clark, M.S., Pegalis, L. and Behan, A. Affective consequences of refusing to help in communal and exchange relationships . Personality and Social Psychology Bulletin , 22, 34–47. (1996) <http://dx.doi.org/10.1177/0146167296221004>

Williamson, G.M. and Schulz, R. Communal orientation, quality of prior relationship and distress among caregivers of Alzheimer's patients Psychology and Aging , 5, 502–507. (1990) <http://dx.doi.org/10.1037/0882-7974.5.4.502>

Williamson, G.M. and Silverman, J.G. Violence against female partners: Direct and interactive effects of family history, communal orientation and peer-related variables . Journal of Social and Personal Relationships , 18, 535–549. (2001) <http://dx.doi.org/10.1177/0265407501184006>

Yoo, S.H., Clark, M.S., Lemay, E.P., Salovey, P. and Monin, J.K. Responding to partners' expression of anger: The role of communal motivation . Personality and Social Psychology Bulletin , 37, 229–241. (2011) <http://dx.doi.org/10.1177/0146167210394205>

<http://dx.doi.org/10.4135/9781446249222.n38>

Handbook of Theories of Social Psychology

Interdependence Theory

Contributors: Paul A.M. Van Lange & Caryl E. Rusult

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Interdependence Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n39>

Print pages: 251-273

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n39>

[p. 251 ↓]

Chapter 39: Interdependence Theory

Abstract

As one of the classic theories of social psychology, interdependence theory has since its earliest formulation (Thibaut and Kelley, 1959) addressed broad classic themes such as dependence and power, rules and norms, as well as coordination and cooperation. Later, Kelley and Thibaut (1978) provided a more comprehensive statement of the theory which allowed researchers to analyze topics such as attribution and self-presentation, trust and distrust, love and commitment, conflict and communication, and risk and self-regulation. Interdependence theory seeks to capture the essence of social life by advancing a conceptual framework for understanding social interaction. In particular, it identifies the most important characteristics of interpersonal situations via a comprehensive analysis of situation structure, and describes the implications of structure for understanding intrapersonal and interpersonal processes (Kelley et al., 2003). Situation structure matters because it is the interpersonal reality within which motives are activated, toward which cognition is oriented, and around which interaction unfolds. This chapter describes key principles of the theory, and illustrates the utility of an interdependence theoretic analysis via a review of phenomena that we may observe everywhere around us – such as regulatory fit, persistence in the face of dissatisfaction, the basis for understanding generosity, and the ebbs and flows of intergroup relations.

Introduction

One of the truly classic theories in the social and behavioral sciences is interdependence theory originally developed by John Thibaut and Harold Kelley in 1959. In the 1998 edition of the *Handbook of Social Psychology*, in his chapter on the historic development of social psychology, Ned Jones made the following prediction

about interdependence theory: "Given the elegance and profundity of this analysis ... there is good reason that its impact will be durable" (1998: 30). Now, more than a decade later, it is clear that interdependence theory has influenced generations of scientists for more than 50 years. It is especially interesting to see that it has stimulated research in various domains of social psychology, including research focusing on within-person processes such as affect and cognition, as well as between-person processes such as behavior and interaction in dyads and groups. Since Thibaut and Kelley (1959) and Kelley and Thibaut (1978), interdependence theoretical concepts and principles have been used to analyze group dynamics, power and dependence, social comparison, [p. 252 ↓] conflict and cooperation, attribution and self-presentation, trust and distrust, emotions, love and commitment, coordination and communication, risk and self-regulation, performance and motivation, social development, and neuroscientific models of social interaction (for recent reviews, see Kelley et al., 2003; Reis, 2008; Rusbult and Van Lange, 2003; Van Lange et al., 2007).

The main focus of interdependence theory is on social interaction, a comprehensive concept that captures the basics of human social life, which helps explain why interdependence theory has been used to understand so many themes for so long. After all, interaction is at the heart of where people live their social lives. Many feelings and emotions are rooted in social interactions, and many beliefs and thoughts are about past or future social interactions. For example, whether a close partner expresses understanding for your bad feelings after you have been mistreated by somebody else is essential for how we feel and think about ourselves, and how we feel and think about the partner – which has strong implications on how we approach a future interaction situation with the partner (and perhaps other people as well). Typically, social interactions exert strong effects in the laboratory, but outside of the lab where interactions often extend over substantial periods of time, social interactions tend to exert even more dramatic effects on us and our relationships. One can indeed go so far as to claim that social interaction colors nearly every phenomenon studied in the social and behavioral sciences, including mental and physical health, personal dispositions, and cognitive and affective experiences (Reis et al., 2000; Rusbult and Van Lange, 2003).

Generally, we argue that the field of psychology would benefit substantially from a social interaction analysis of human psychology, and suggest that interdependence

theory can play an important role in this respect. Interdependence theory is one of the few social psychological theories that provide a comprehensive analysis with a strong orientation toward conceptualizing *interpersonal* structure and processes (Kelley and Thibaut, 1978; Kelley et al., 2003; Thibaut and Kelley, 1959). Analogous to contemporary physics – where the relations between particles are as meaningful as the particles themselves – in interdependence theory, *between-person* relations are as meaningful as the individuals themselves (Rusbult and Van Lange, 2003). Indeed, concepts such as coordination, trust, cooperation, communication, commitment can only be understood in terms of social interaction, and many of the needs, motives, and processes that receive considerable attention in contemporary social psychology – such as need-to-belong, uncertainty-management, self-regulation – are often oriented in the service of dealing with the threats and opportunities of social interaction.

In this chapter, we outline the key principles of interdependence theory, provide a historical account of its roots and development over the five decades, and outline some prospects for the future. In doing so, we also provide a narrative of major challenges that the founders of interdependence theory (must have) faced, and those that the next generation, along with Hal Kelley, have faced. It will also become clear that interdependence theory is growing while benefiting from the solid foundation (and more) that the fathers of interdependence theory have provided. We conclude by describing broad implications for various social psychological phenomena and applications in several societal domains.

Interdependence Structure

Interdependence theory uses two formal tools to represent the outcomes of interaction–matrices and transition lists (Kelley, 1984; Kelley and Thibaut, 1978). The purpose of these formal representations is to precisely specify the character of situation structure – to describe the ways in which people can affect one another's outcomes during the course of interaction. *Interaction* describes two people's (*A* and *B*) needs, thoughts, [p. 253 ↓] motives, and behaviors in relation to one another in the context of the specific interdependence situation (*S*) in which their interaction transpires (Kelley et al., 2003). Expressed formally, $I = f(S, A, B)$. To predict what will transpire in an interaction between two persons, we must consider (a) what situation they confront

(e.g., are their interests at odds, does one hold greater power?), (b) person A's needs, thoughts, and motives with respect to this interaction (i.e., which traits or values are activated, how does he feel about person B?), and (c) person B's needs, thoughts, and motives with respect to this interaction. In the following, we replace persons A and B with John and Mary, two names that have often been used to illustrate the formal logic of interdependence theory. The model involving the situation and the two persons is sometimes referred to as the SABI model, an acronym for Situation, persons A and B that collectively account for Interaction (e.g., Holmes, 2002; Kelley et al., 2003; Van Lange et al., 2007; see the principle of structure, and the principle of interaction, [Box 39.1](#)).

The precise outcomes of an interaction – the degree to which John and Mary experience it as satisfying – depend on whether the interaction gratifies (versus frustrates) important needs, such as security, belongingness, and exploration (cf. Baumeister and Leary, 1995; Fiske, 2004). Interaction not only yields *concrete outcomes*, or immediate experiences of pleasure versus displeasure, but also *symbolic outcomes*, or experiences that rest on the broader implications of interaction (e.g., Rusbult and Van Lange, 1996). For example, if John and Mary disagree about where to dine yet John suggests Mary's favorite restaurant, Mary not only enjoys the concrete benefits of good food and wine, but also enjoys the symbolic pleasure of perceiving that John is responsive to her needs.

By analyzing how each person's possible behaviors would affect each person's outcomes, we can discern the structure of a situation with respect to degree and type of dependence, examining: (a) *actor control* – the impact of each person's actions on his or her own outcomes; (b) *partner control* – the impact of each person's actions on the partner's outcomes; and (c) *joint control* – the impact of the partners' joint actions on each person's outcomes. And by examining the across-cell association between outcomes, we can discern *covariation of interests*, or the extent to which the partners' outcomes are correlated. These components define four structural dimensions; two additional dimensions have also been identified more recently (all six are described below; Kelley et al., 2003). Most situations are defined by their properties with respect to two or more dimensions. For example, the *prisoner's dilemma*, *hero*, and *chicken* situations all involve moderate and mutual dependence along with moderately conflicting interests, but these neighboring situations also differ in the magnitude

of actor control, partner control, and joint control, as well as in their implications for interaction.

All conceivable combinations of the six properties define a very large number of patterns. However, we can identify at least 20 to 25 prototypes (Kelley et al., 2003). Everyday situations resemble these abstract patterns, sharing common interpersonal problems and opportunities. For example, the *twists of fate* situation is one wherein each partner, at some point, might unexpectedly find himself or herself in a position of extreme unilateral dependence; this sort of situation is characteristic of health crises and other reversals of fortune. And as another example, the *prisoner's dilemma* is a situation wherein each person's outcomes are more powerfully influenced by the partner's actions than by his or her own actions; this sort of situation is characteristic of interactions involving mutual sacrifice, trading favors, and free-riding. Everyday situations that share the same abstract pattern have parallel implications for motivation, cognition, and interaction.

Importance of Interdependence Structure

Why should we care about interdependence structure? To begin with, structure in itself [p. 254 ↓] reliably influences behavior. For example, situations with structure resembling the *threat* situation reliably yield demand-withdraw patterns of interaction – demands for change on the part of the lower power actor, met by withdrawal and avoidance on the part of the higher power partner (Holmes and Murray, 1996). And situations with structure resembling the *chicken* situation reliably yield interaction centering on establishing dominance and sustaining one's reputation (Nisbett and Cohen, 1996). In short, the structure of situations often directly shapes behavior above and beyond the specific goals and motives of interacting individuals.

Box 39.1 Overview of Basic Assumptions of Interdependence Theory

The Principle of Structure (“The Situation”)

Understanding interdependence features of a situation are essential to understand psychological process (motives, cognition, and affect), behavior, and social interaction. The features are formalized in a taxonomy of situations, which are degree of dependence, mutuality of dependence, covariation of interest, basis of dependence, temporal structure, and information availability.

The Principle of Transformation (What People Make of “The Situation”)

Interaction situations may be subject to transformations by which individualist consider consequences of own (and other's) behavior in terms of outcomes for self and others and in terms of immediate and future consequences. Transformation is a psychological process that is guided by interaction goals, which may be accompanied and supported by affective, cognitive, and motivational processes.

The Principle of Interaction: SABI: $I = f(A, B, S)$

Interaction is a function of two persons (persons *A* and *B*) and (objective properties) of the situation. The situation may activate particular motives, cognitive, and affective experiences in persons *A* and *B*, which ultimately through their mutual responses in behavior yield a particular pattern of interaction.

The Principle of Adaptation

Repeated social interaction experiences yield adaptations that are reflected in relatively stable orientations to adopt particular transformations. These adaptations are probabilistic and reflect (a) differences in orientation between people across partners and situations (dispositions), (b) orientations that people adopt to a specific interaction partner (relationship-specific orientations), and (c) rule-based inclinations that are shared by many people within a culture to respond to a particular classes of situation in a specific manner (social norms).

Moreover, specific structural patterns present specific sorts of problems and opportunities, and therefore (a) logically imply the relevance of specific goals and motives, and (b) permit the expression of those goals and motives. The term *affordance* nicely describes what a situation makes possible or may activate (see [Table 39.1](#), which provides an overview of possible affordances). For example, situations with uncertain information afford misunderstanding, and invite reliance on generalized schemas regarding partners and situations; generalized schemas carry less weight when information is more complete. In short, situation structure matters because it is the interpersonal reality within which motives are activated, toward which cognition is oriented, and around which interaction unfolds.

Dimensions of Interdependence Structure

Level of dependence describes the degree to which an actor relies on an interaction partner, in that his or her outcomes are influenced by the partner's actions. If Mary can obtain good outcomes irrespective of John's actions (high actor control), she is independent; she is dependent to the extent that John can (a) unilaterally determine her pleasure versus [p. 255 ↓] displeasure (partner control) or (b) in combination with Mary's actions determine her pleasure versus displeasure (joint control). Increasing dependence tends to cause increased attention to situations and partners, more careful and differentiated cognitive activity, and perseverance in interaction (e.g., Fiske, 1993; Rusbult, 1983). As noted in [Table 39.1](#), dependence affords thoughts and motives

centering on comfort versus discomfort with dependence and independence. For example, high dependence situations will activate Mary's trait-based reluctance to rely on others, her discomfort with dependence will strongly shape her behavior, and her discomfort will be particularly evident to others; in low dependence situations, this trait will be less visible and less relevant for her behavior.

Table 39.1 The six dimensions of situational structure and their affordances (after Holmes, 2002; and Kelley et al., 2003)

<i>Situation Dimension</i>	<i>Relevant Motives</i>
1 Level of dependence	Comfort versus discomfort with dependence; comfort versus discomfort with independence
2 Mutuality of dependence	Comfort versus discomfort with vulnerability (as dependent) Comfort versus discomfort with responsibility (as power holder)
3 Basis of dependence	Dominance (leading) versus submissiveness (following) Assertiveness versus passivity
4 Covariation of interests	Prosocial versus self-interested motives (rules for self) Trust versus distrust of partner motives (expectations about others)
5 Temporal structure	Dependability versus unreliability Loyalty versus disloyalty
6 Information availability	Openness versus need for certainty Optimism versus pessimism

Mutuality of dependence describes whether two people are equally dependent upon each other. Nonmutual dependence entails differential power – when Mary is more dependent, John holds greater power. The less dependent partner tends to exert greater control over decisions and resources, whereas the more dependent partner carries the greater burden of interaction costs (sacrifice, accommodation) and is more vulnerable to possible abandonment; threats and coercion are possible (e.g., Attridge et al., 1995; Murray et al., 2006). Interactions with mutual dependence tend to feel “safer” and are more stable and affectively serene (less anxiety, guilt). Situations with nonmutual dependence afford the expression of comfort versus discomfort with another having control over your outcomes (e.g., feelings of vulnerability, for the dependent partner) along with comfort versus discomfort with you having control over other's outcomes (e.g., feelings of responsibility, for the powerful partner; see [Table 39.1](#)). For example, unilateral dependence will activate John's insecurity, and his insecurity will powerfully shape his behavior and be highly visible to others; in mutual dependence situations his insecurity will be less visible and less relevant to predicting his behavior.

Basis of dependence describes precisely how partners influence one another's outcomes – the relative importance of partner versus joint control as source of

dependence. With partner control, the actor's outcomes rest in the partner's hands, so interaction often involves promises or threats as well as the activation of morality norms ("This is how decent people behave"); common interaction patterns may include unilateral action (when partner control is nonmutual) or tit-for-tat or turn-taking (when partner control is mutual; for example, Clark et al., 1998; Fiske, 1992). In contrast, joint control entails contingency-based coordination of action, such that ability-relevant traits become more important, including intelligence, [p. 256 ↓] initiative-taking, and strategic skills; rules of conventional behavior carry more sway than morality norms ("This is the normal way to behave;" for example, Finkel et al., 2006; Turiel, 1983). Basis of dependence affords the expression of dominance versus submissiveness and assertiveness versus passivity, as well as skill such as social intelligence (see [Table 39.1](#)).

Covariation of interests describes whether partners' outcomes correspond versus conflict – whether partners' joint activities yield similarly gratifying outcomes for John and Mary. Covariation ranges from perfectly corresponding patterns through mixed motive patterns to perfectly conflicting patterns (zero-sum). Given corresponding interests, interaction is easy – John and Mary simply pursue their own interests, simultaneously producing good outcomes for the other. In contrast, situations with conflicting interests tend to generate negative cognition and emotion (greed, fear) and yield more active and differentiated information-seeking and self-presentation ("Can Mary be trusted?;" for example, Surra and Longstreth, 1990; Van Lange et al., 1997). Situations with conflicting interests afford the expression of cooperation versus competition and trust versus mistrust (see [Table 39.1](#)) – in such situations, John may demonstrate his prosocial motives as well as his trust in Mary.

Temporal structure is a fifth important structural dimension – one that captures dynamic and sequential processes. As a result of interaction, some future behaviors, outcomes, or situations may be made available and others may be eliminated. John and Mary may be passively moved from one situation to another or they may be active agents in seeking such movement. *Extended situations* involve a series of steps prior to reaching a goal (e.g., investments leading to a desirable outcome). *Situation selection* describes movement from one situation to another, bringing partners to a situation that differs in terms of behavioral options or outcomes – for example, Mary may seek situations entailing lesser interdependence, or John may confront the juncture between a present

relationship and an alternative relationship by derogating tempting alternatives (e.g., Collins and Feeney, 2004; Miller, 1997). Temporally extended situations afford the expression of self-control, delay of gratification, and the inclination to “stick with it” – dependability versus unreliability, as well as loyalty versus disloyalty (e.g., Mischel, [Chapter 26](#), this volume) (see [Table 39.1](#)).

Information availability is the final structural dimension: Do John and Mary possess certain versus uncertain information about: (a) the impact of each person's actions on each person's outcomes; (b) the goals and motives guiding each person's actions; and (c) the opportunities that will be made available versus eliminated as a consequence of their actions? Certain information is critical in novel or risky situations and in interactions with unfamiliar partners. Accordingly, partners engage in a good deal of information exchange during the course of interaction, engaging in attributional activity to understand one another and the situation (e.g., Collins and Miller, 1994). People may also use representations of prior interaction partners to “fill in the informational gaps” in interaction with new partners, or may develop frozen expectations that reliably color their perceptions of situations and partners (e.g., Andersen and Chen, 2002; Holmes, 2002, 2004). For example, people may generally rely on the belief that the most people are (rationally) self-interested, which in turn may help them to fill in the blanks when faced with incomplete information about another person's actions (Vuolevi and Van Lange, 2010). As another example, people with avoidant attachment may perceive a wide range of situations as risky, anticipate that partners are likely to be unresponsive, and readily forecast problematic interactions. Thus, uncertain information affords, among other things, the expression of openness versus the need for certainty, as well as optimism versus pessimism (see [Table 39.1](#)).

[p. 257 ↓]

Interdependence Processes

Recall that interaction ($I = f[S, A, B]$) is shaped not only by interdependence structure (S), but also by partners' needs, thoughts, and motives in relation to one another (A and B) in the context of the situation in which their interaction unfolds (SABI, see Principle of Interaction, [Box 39.1](#)). Thus, we must add to our structural analysis a complementary

analysis that describes how John and Mary react to the situations they encounter. How do they psychologically transform specific situations, responding on the basis of considerations other than tangible self-interest? What role do mental events and habits play in shaping this process, and how do partners seek to understand and predict one another? And how do people develop relatively stable tendencies to react to specific situations in specific ways?

Transformation Process

To describe how situation structure affects motivation, interdependence theory distinguishes between: (a) the *given situation* – preferences based on self-interest (the “virtual structure” of a situation); and (b) the effective situation – preferences based on broader considerations, including concern with the partner's interests, long-term goals, or strategic considerations (Kelley and Thibaut, 1978; Van Lange and Joireman, 2008). Psychological *transformation* describes the shift in motivation from given to effective preferences. People typically behave on the basis of transformed preferences – considerations other than immediate self-interest guide our actions. But people sometimes behave on the basis of given preferences; this is likely in simple situations for which no broader considerations are relevant, when people lack the inclination or wherewithal to take broader considerations into account, and in situations involving time pressure or constrained cognitive capacity.

Transformations are often conceptualized decision rules that a person (often implicitly) adopts during interaction (Kelley et al., 2003; Murray and Holmes, 2009; Van Lange et al., 2007, see the principle of transformation, [Box 39.1](#)). People may follow rules that involve sequential or temporal considerations, such as waiting to see how the partner behaves or adopting strategies such as tit-for-tat or turn-taking. Other rules reflect differential concern for one's own and a partner's outcomes, including altruism, or maximizing the partner's outcomes; cooperation, or maximizing combined outcomes; competition, or maximizing the relative difference between one's own and the partner's outcomes; and individualism, or maximizing one's own outcomes irrespective of the partner's outcomes.

Transformation is particularly visible when a given situation structure dictates one type of behavior yet personal traits or values dictate another type of behavior. When people act on the basis of transformed preferences, we are able to discern their personal traits and motives. For example, when Mary helps John with yard work rather than going out with her friends, she communicates concern for his welfare. The transformation process is thus the point at which the “rubber meets the road,” or the point at which *intrapersonal* processes – cognition, affect, and motivation – operate on specific situations in such a manner as to reveal the unique self.

Cognition, Affect, and Habit

Human intelligence is interpersonal – cognitively and affectively, we are well prepared to construe the world in terms of interdependence (Rusbult and Van Lange, 2003). Mental events are geared toward discerning what a situation is “about,” evaluating that structure in terms of one's own needs and motives, perceiving the partner's needs and predicting his or her motives, and forecasting implications for future interactions (e.g., Kelley, 1984). Situation structure partially shapes cognition and affect. For example, the prisoner's dilemma entails a choice between benefiting the partner at low cost to the self versus benefiting the self at substantial cost [p. 258 ↓] to the partner. The characteristic blend of fear and greed that is afforded by this situation serves as a rather automatic indicator of the essential opportunities and constraints of this type of situation.

The transformation process is often driven by the cognition and affect that a situation affords. For example, Mary is likely to exhibit self-centered or antisocial transformation when she experiences greedy thoughts and desires (“It'd be nice to take a free ride”) or feels fearful about John's motives (“Will he exploit me?”). Cognition and emotion are also shaped by distal causes – by the values, goals, and dispositions that are afforded by the situation. For example, Mary's reaction to situations with conflicting interests will be colored by the value she places on fairness, loyalty, or communal norms (versus greed), as well as by whether she trusts John (or alternatively, fears him). Thus, the mental events that underlie transformation are functionally adapted to situation structure, and take forms that are relevant to that structure.

At the same time, the transformation process does not necessarily rest on extensive mental activity. As a consequence of adaptation to repeatedly encountered patterns, people develop habitual tendencies to react to specific situations in specific ways, such that transformation often transpires with little or no conscious thought (e.g., Rusbult and Van Lange, 1996). For example, following repeated interaction in situations with prisoner's dilemma structure, John and Mary may automatically exhibit mutual cooperation, with little or no cognition or affect. Mediation by mental events is more probable in novel situations with unknown implications, in risky situations with the potential for harm, and in interactions with unfamiliar partners.

Communication, Attribution, and Self-Presentation

During the course of interaction, partners convey their goals, values, and dispositions using both direct and indirect means. Communication entails self-presentation on the part of one person and attribution on the part of the other. As noted earlier, the material for self-presentation and attribution resides in the disparity between the given and effective situations, in that deviations from self-interested behavior reveal an actor's goals and motives (e.g., Rusbult and Van Lange, 2003). Thus, the ability to communicate self-relevant information is limited by interdependence structure – that is, specific situations afford the display of specific motives. For example, it is difficult for people to convey trustworthiness (or to discern it) in situations with correspondent interests; in such situations, “trustworthy” behavior aligns with “self-interested” behavior.

People engage in *attributional activity* to understand the implications of a partner's actions, seeking to predict future behavior and to explain prior behavior in terms of situation structure versus underlying dispositions. Expectations are not particularly accurate in interactions with new partners, in that they must be based on probabilistic assumptions about how the average person would react in a given situation; in longer-term relationships, expectations can also be based on knowledge of how a partner has behaved across a variety of situations. And *self-presentation* describes people's attempts to communicate their motives and dispositions to one another. Of course, self-presentation may sometimes be geared toward concealing one's true preferences

and motives. Moreover, given that people do not always hold complete information about their partners' given outcomes, they may sometimes mistakenly assume that a partner's behavior reflects situation structure rather than psychological transformation. For example, Mary's loyalty or sacrifice may not be visible if John fails to recognize the costs she incurred.

Adaptation

When people initially encounter specific situations, the problems and opportunities [p. 259 ↓] inherent in the situation will often be unclear. In such novel situations, Mary may systematically analyze the situation and actively reach a decision about how to behave, or she may simply react on the basis of impulse. Either way, experience is acquired. If her choice yields good outcomes, she will react similarly to future situations with parallel structure; if her choice yields poor outcomes, she will modify her behavior in future situations with parallel structure. *Adaptation* describes the process by which repeated experience in situations with similar structure gives rise to habitual response tendencies that on average yield good outcomes. Adaptations may be embodied in interpersonal dispositions, relationship-specific motives, or social norms (Rusbult and Van Lange, 1996, see the principle of adaptation, [Box 39.1](#)).

Interpersonal dispositions are actor-specific inclinations to respond to particular classes of situation in a specific manner across diverse partners (Kelley, 1983). Dispositions emerge because over the course of development, different people experience different histories with different partners, confronting different sorts of interaction opportunities and problems. As a result of adaptation, John and Mary acquire dispositional tendencies to perceive situations and partners in specific ways, and specific sorts of transformations come to guide their behavior. Thus, the "self" is the sum of one's adaptations to previous situations and partners (such adaptations are determined also by needs and motives that are biologically based). For example, if John's mother employed her power in a benevolent manner, gratifying his childhood needs and serving as a secure base from which he could explore, John will have developed trusting and secure expectations about dependence (for a review, see Fraley and Shaver, 2000).

Relationship-specific motives are inclinations to respond to particular classes of situation in a specific manner with a specific partner (Rusbult and Van Lange, 2003). For example, commitment emerges as a result of dependence on a partner, and is strengthened by high satisfaction (John gratifies Mary's most important needs), poor alternatives (Mary's needs could not be gratified independent of her relationship), and high investments (important resources are bound to her relationship). Commitment colors emotional reactions to interaction (feeling affection rather than anger) and gives rise to habits of thought that support sustained involvement (use of plural pronouns; for example, Agnew et al., 1998). In turn, benevolent thoughts encourage prosocial transformation. For example, strong commitment promotes prosocial acts such as sacrifice, accommodation, and forgiveness (e.g., Finkel et al., 2002; Rusbult et al., 1991; Van Lange et al., 1997).

Social norms are rule-based, socially transmitted inclinations to respond to particular classes of situation in a specific manner (Thibaut and Kelley, 1959). For example, most societies develop rules regarding acceptable behavior in specific types of situation; rules of civility and etiquette regulate behavior in such a manner as to yield harmonious interaction. Partners frequently follow agreed-upon rules regarding resource allocation, such as equity, equality, or need (Deutsch, 1975). Such rules may govern a wide range of interactions or may be relationship-specific (e.g., communal norms in close relationships; Clark et al., 1998; Fiske, 1992). Norms not only govern behavior, but also shape cognitive experiences. For example, in interactions guided by communal norms, partners neither monitor nor encode the extent of each person's (short-term) contributions to the other's welfare.

Development of Interdependence Theory: An Interpersonal Account

As noted earlier, the history of interdependence theory is strongly shaped by the longstanding collaboration and friendship between Harold Kelley and John Thibaut. A sketchy summary of the history of interdependence [p. 260 ↓] theory is provided in Table 39.2. Our narrative is written from the perspective of Harold Kelley, as the authors of this chapter interacted much more with "Hal" (1921–2003) than with John

Thibaut (1917–1986), which is why we refer to the former as Hal and the latter as John Thibaut. The collaboration between Thibaut and Kelley started when Hal was invited to write a chapter on “group problem solving” for the *Handbook of Social Psychology*. Hal invited John Thibaut, whom he knew well from the Research Center for Group Dynamics at MIT, to collaborate on writing this chapter. This decision, so he described informally, was one of the very best in his academic career. There was an interpersonal fit from the very beginning, and they wrote a beautiful chapter, inspired by some of the notions put forward by Kurt Lewin, in which they analyzed the interdependence between individuals in their pursuit of group goals. The major themes – interdependence and social interaction – were discussed in a manner that was predictive of [p. 261 ↓] their later collaboration, one that lasted for three decades until the death of John Thibaut in 1986. They developed a collaboration that was characterized by many travels between Malibu and Chapel Hill, by deep friendship and tremendous mutual respect, by equality (they were both follower and leader) as well as by similarity and complementarity. To magnify the latter (for illustration purposes), the natural distribution of tasks was that Hal focused more strongly on analysis of situations, while John Thibaut focused more strongly on connections with the various literatures inside and outside of psychology. They were also complementary in that Hal's interests focused more on the dyad (later relationships) whereas John's interests' focused more on the (small) group.

Table 39.2 Brief historical overview of interdependence theory

1959	Thibaut, J.W. and Kelley, H.H. (1959) <i>The Social Psychology of Groups</i> . New York: Wiley.	Provides social exchange analysis of interactions and relationships individuals in dyads and small groups. Uses games as a conceptual tool and focuses on analysis of dependence, power, rewards, costs, needs, and outcomes in exchange relations. Introduces new concepts such as comparison level and comparison level of alternatives (CL and CL-alt) to understand relationship satisfaction and stability.
1978	Kelley, H.H., and Thibaut, J.W. (1978) <i>Interpersonal Relations: A Theory of Interdependence</i> . New York: Wiley.	Provides comprehensive analysis of interaction situations in terms of four dimensions, labeled as degree of dependence, mutuality of dependence, correspondence of outcomes, and basis of dependence. Introduces transformation from given to effective matrix, thereby formalizing broader interaction goals than immediate self-interest. Adopts a functional analysis of transformations, thereby recognizing social learning of transformation rules, and its functional value for particular domains of situation.
2003	Kelley, H.H. et al. (2003) <i>An Atlas of Interpersonal Situations</i> . New York: Cambridge University Press	Provides an overview of 21 basic interaction situations, which are analyzed in terms of interdependence features, the psychological processes that they afford, and the interaction processes that they might evoke. Extends the taxonomy of situations by two additional dimensions to yield six dimensions, including (a) degree of dependence, (b) mutuality of dependence, (c) basis of dependence, (d) covariation of interest (was formerly referred to as correspondence of outcomes), (e) temporal structure, and (f) information availability.
At present and in the future: Integrates interdependence theory with principles of evolutionary theory to understand adaptation as a function of the situational structure. Extends interdependence theory to neuroscientific models of the social mind. Re-extends interdependence theory to group processes and relationships between groups.		

They then wrote a book (Thibaut and Kelley, 1959) that was inspired by social exchange theory (in particular, Homans, 1950) and by game theory and decision theory (in particular, a highly influential book by Luce and Raiffa, 1957). Essentially, they analyzed social interactions in dyads and small groups in terms of patterns of social exchange, thereby using games as the conceptual tool – to be able to delineate the patterns of interdependence, such as rewards and costs, and power and dependence. They also introduced new concepts such as comparison level and comparison level for alternatives (CL and CL-alt) to provide a strong conceptual analysis for the differences between satisfaction and dependence. This book was a great success and a must-read for any social psychologist at that (or any) time (see Jones, 1998).

After nearly two decades, Kelley and Thibaut (1978) modestly expressed the belief that their new analysis – an interdependence analysis – might well reach the standards of a theory. While the origins were captured in the 1959 book, interdependence theory was now formally born (Hal and John were careful scientists and they would reserve the label theory only for those kinds of conceptual analysis that would pass stringent tests of scientific rigor – probably defined by Hal and John Thibaut in terms of clear logic

and wide breadth of relevance). In that book, they presented interdependence theory, and it became immediately clear that many years were devoted to very basic theoretical issues.

One decision they faced was whether behavior was primarily based on the given matrix (i.e., on the basis of immediate self-interest) or whether the theory should be extended to include broader considerations. Informed by research during the sixties and seventies, they agreed on the latter and provided a logical framework for a number of fundamental transformations, which they labeled as MaxJoint (enhancement of joint outcomes), MinDiff (minimization of absolute differences in outcomes for self and others), MaxRel (maximization of relative advantage over other's outcomes), and the like. These transformations were also inspired by the work of Messick and McClintock, and many others around the globe, who had already provided empirical evidence for some transformations in their research using experimental games as empirical tools (e.g., Messick and McClintock, 1968). Hal and John also outlined other types of transformation, which emphasize the idea that people respond to contingencies and expected implications of present behavior for the future. Another key difference with the earlier book was that it emphasized the functional value of various transformations. In short, this book contributes logic to the question, What do "people make of situations?" (see also Kelley et al., 2003).

Thus, the classic Kelley and Thibaut interdependence analysis became a comprehensive theory encompassing (a) a formal analysis of the "objective" properties of a situation with the help of a taxonomy of situations, (b) a conceptualization of psychological process in terms of transformations, including motives, cognition, and affect (what do people make of the situation?), and (c) behavior and social interaction – which resulted from both the objective properties of the situation and what both persons made of it. Moreover, they emphasized (d) adaptation and learning, as longer-term orientations that [p. 262 ↓] may grow out of experience. Inspired by the work of Messick and McClintock (1968), and their own (Kelley and Stahelski, 1970), they also suggested that people might differ in their "transformational tendencies." These adaptations were later conceptualized in terms of dispositions, relationship-specific motives, and social norms (see Rusbult and Van Lange, 1996).

Over time, numerous people were inspired by the “logic” of interdependence theory – its assumptions, the reasoning, and last but not least, its focus. Logic is one thing, but it appeared to have considerable breadth. And so researchers in areas as diverse as altruism, attribution, coordination, conflict, cooperation, competition, delay of gratification, exchange, investments, fairness, justice, love, power, prosocial behavior, trust, sacrifice, self-presentation, stereotyping, hostility, and aggression in the context of dyads, ongoing relationships (close or not), and groups (small and larger, ongoing or not) either found it exceptionally useful or were inspired by it. Also, researchers studying environmental issues, organizational issues, and political issues have fruitfully used principles from interdependence theory (for a comprehensive review, see Rusbult and Van Lange, 1996, 2003; Van Lange and Joireman, 2008). The list of authors is too long to summarize here, but we wish to note that if we were to list them, it would become clear that interdependence theory had a strong influence in various countries even in the pre-Internet era (most notably, Austria, Australia, Belgium, Canada, France, Germany, Italy, Japan, the Netherlands, New Zealand, Poland, Sweden, the United Kingdom, and the US), and that influence spanned successive generations, so that it is fair to say that it has strong appeal to young-, mid-, and late-career scientists.

To illustrate from the experience of the Atlas project group, John Holmes had worked with John Thibaut and used principles of interdependence theory in his work on trust and conflict (as well as on motivation-management in relationships; Holmes and Rempel, 1989; Murray and Holmes, 2009). Caryl Rusbult developed the investment model of commitment processes, a framework that was deeply rooted in interdependence theoretic principles, to understand persistence and commitment processes in ongoing relationships (see Rusbult and Van Lange, 2003; Rusbult et al., 2006). Paul Van Lange was intrigued by Kelley and Thibaut's taxonomy of situations (“structure”) and transformations (what “persons” make of situations) and found it very useful for his research on social value orientation as well as for understanding the functionality of generosity in social dilemmas (see Van Lange et al., 1997, 2002). Norbert Kerr found an interdependence perspective useful for understanding group-related issues as diverse as motivation and performance, cooperation, and free-riding in social dilemmas (see Baron and Kerr, 2003; Kerr and Tindale, 2004). It was Harry Reis who not only had used interdependence theory in his research on intimacy and responsiveness in relationships (e.g., Reis, 2008; Reis et al., 2000), but also had

the vision and skills in getting this group of people together at a joint meeting of the Society of Experimental Social Psychology (SESP) and the European Association of Experimental Social Psychology (EAESP) in Washington in 1995. This resulted in a six-year collaboration and eventually the publication of the *Atlas of Interpersonal Situations* (Kelley et al., 2003).

The group came together at various meetings, often right before or after a major social psychological conference in Europe or the US. There were two meetings that were independent of a conference. First, in 1996, during Caryl's sabbatical at the VU University in Amsterdam, we held a series of eight-hour (nearly nonstop) daily sessions for about seven days. Hal, Caryl Rusult, and Paul Van Lange discussed aspects of what was later called "temporal structure" and drafted an outline for chapters for the book. But fortunately, those not present later corrected a tendency to embrace complexity rather than parsimony. Second, the other series of meetings was held in 2000 in Boca Raton, Florida, [p. 263 ↓] generously sponsored by Bibb Latané. At this series of meetings, we discussed the various drafts of the chapters and reached final consensus over the situations that should – or should not – be included in the book.

The *Atlas* by Kelley et al. (2003) extended Kelley and Thibaut (1978) in very important ways, but perhaps most notably by analyzing 21 situations and by adding two dimensions to the four dimensions of interdependence that Kelley and Thibaut already had previously identified. The added dimensions were (a) temporal structure and (b) information availability. The first copy of the book was published ahead of schedule (thanks to Harry Reis and our publisher, Cambridge University Press) and was given to Hal Kelley about a week or less before he died. Caryl, John, Norb, and Paul saw the first copy at the Society for Personality and Social Psychology (SPSP) meeting in Los Angeles, in February 2003, shortly before a memorial service for Hal at UCLA. Also, Hal suggested earlier that we dedicate the book to the memory of John Thibaut, and that suggestion received strong support, in synchrony, from us all.

As one of us (PvL) edits this chapter, I might be indulged in saying a bit more about the contributions to the development of interdependence theory of my late colleague, collaborator, and dear friend, Caryl Rusult. Needless to say, Caryl Rusult was a major contributor to the *Atlas* throughout all six years. As a UCLA undergraduate (where Hal was professor) and UNC (Chapel Hill) graduate student (where John Thibaut

was professor) and later faculty member, she developed a strong commitment to interdependence theory. It was only two days before she died (far too young at the age of 57 in January 2010) that she and I re-evaluated the various projects we had worked on together. We decided that the comprehensive review (Rusbult and Van Lange, 1996) and the *Atlas* joint venture (Kelley et al., 2003) were among the highlights of our long-standing collaboration. We truly enjoyed talking about interdependence theory – its logic, the ways in which it needs to be communicated and extended, and its implications for basic questions about relationship processes (Caryl's passion) and human cooperation (Paul's focus). We also frequently discussed “applications” of interdependence theory by examining why and when an interdependence-theoretical analysis mattered. This is the question that we address next.

Applications of Interdependence Theory

To comprehend the utility of interdependence concepts it is important to “see them in action” – to perceive the theoretical, empirical, and societal benefits of these concepts in advancing our understanding of specific psychological phenomena. In particular, we suggest that interdependence theory is especially useful for understanding relationship persistence and stability, interpersonal generosity, as well as other broad topics – such as goal pursuits in relationships, and understanding of group processes.

Understanding Goal Pursuits

Our first example illustrates a simple point: *interdependence matters*. In fact, interdependence shapes many psychological processes that might seem to be thoroughly actor-based and *intrapersonal*, such as individual goal pursuits. Goals are end states that give direction to behavior, either as overarching life plans or as simple everyday endeavors. Traditional models of goal pursuit have employed *intrapersonal* explanations, examining individual-level processes such as goal-plan directed behavior, self-regulation, or goal-behavior disparities (e.g., Carver and Scheier, 1998; Mischel, [Chapter 26](#), this volume). The success of goal pursuit has been argued to rest on actor-level variables such as goals, traits, skills, and motivation. A notable approach in

this tradition is regulatory focus and regulatory fit theories which suggest that people are [p. 264 ↓] more likely to achieve goals when they approach them in a manner that fits their regulatory orientation – when they approach promotion-ideal self-goals of accomplishment in an eager manner and approach prevention-ought self-goals of security in a vigilant manner (Higgins, 1997, 2000, 2011).

An interdependence analysis shares some of these assumptions, but extends them in interesting directions. Indeed, research using diverse empirical techniques has revealed that in ongoing relationships, people enjoy greater movement toward their ideal selves not only when (a) they, themselves, possess strong promotion orientation (actor control), but also when (b) their partners possess strong promotion orientation (partner control) (parallel negative associations are evident for prevention orientation; Righetti and Rusult, 2007). Indeed, partners with a strong promotion orientation support the actor's movement toward the ideal self because such partners more reliably elicit key components of the actor's ideal-related eagerness. Some empirical support was also obtained for a third form of fit: Above and beyond the above-noted actor and partner effects, there is some evidence for a joint control effect, such that (c) actor–partner commonality in regulatory orientation also influences each person's movement toward the ideal self. Thus, the fact that goal pursuit and attainment are powerfully and reliably influenced by interdependence processes suggests that there is much to recommend in an interdependence theoretic analysis. Interdependence matters.

Understanding Persistence

Our second example illustrates the fact that *interdependence structure* matters. Indeed, structure can often help explain otherwise inexplicable phenomena, such as why attitudes do not always satisfactorily predict behavior, or why people sometimes persist in situations that are not particularly satisfying. Traditionally, persistence has been explained by reference to positive affect: people persevere in specific endeavors because they have positive explicit or implicit attitudes about the endeavor; people persevere in specific jobs or relationships because they feel satisfied with them (e.g., Ajzen, 1991; Greenwald et al., 1998). The affect construct has been operationally defined in terms of satisfaction level, positive attitudes, liking, or attraction.

An important challenge to this “feel good” model of persistence (“So long as it feels good, I’ll stick with it”) is to be found in situations wherein people persevere despite the existence of negative affect. Clearly, people sometimes persevere even though they hold negative attitudes about behavior-relevant attitude objects; people sometimes stick with jobs or marriages despite feelings of dissatisfaction. Persistence in an abusive relationship is a particularly telling illustration: surely people do not persist because they are delighted with such relationships. Some authors have sought to account for such inexplicable persistence in terms of trait-based explanations – by reference to a victim’s low self-esteem or learned helplessness (e.g., Aguilar and Nightingale, 1994; Walker, 2000). Inexplicable persistence is thus assumed to be an actor effect – people persevere because of something peculiar or unhealthy about themselves.

In contrast, an interdependence analysis explains persistence more broadly, by reference to the nature of an actor’s dependence. To the extent that people are more dependent upon their jobs or relationships, they are more likely to persist in them; the greater their dependence upon a distal goal, the more likely they are to persist in pursuit of the goal. In relationships, dependence is strengthened by increasing satisfaction (are important needs gratified?), declining alternatives (could important needs be gratified elsewhere?), and increasing investments (are important resources linked to the line of action?; see Rusbult et al., 2006). For example, Mary may persevere in an abusive relationship not necessarily because she has low self-esteem or has acquired a pattern of learned helplessness, but rather, for reasons resting on structural [p. 265 ↓] dependence – because she is heavily invested in remaining with her partner (e.g., she is married to John or has young children with him) or has poor alternatives (e.g., she has no driver’s license or possesses poor employment opportunities; Rusbult and Martz, 1995).

Why should scientists favor an interdependence-based analysis of persistence? For one thing, positive affect is not particularly reliable – affect ebbs and flows even in the most satisfying jobs and relationships, such that “feeling good” is not sufficient to sustain long-term persistence. In addition, actor-based explanations would appear to be limited in light of clear evidence for dependence-based causes of persistence (e.g., Mary may have invested too much to quit). Moreover, interdependence-based explanations imply unique intervention strategies. For example, if we seek to enhance Mary’s freedom to persist versus cease involvement with John, an actor-based explanation might favor

psychotherapy geared toward raising self-esteem or eliminating learned helplessness. In contrast, an interdependence-based explanation might inspire interventions designed to reduce (unilateral) dependence – for example, improving the quality of Mary's economic alternatives via driving lessons or job training. Also, in therapy, the focus may not only be on some fluctuation in satisfaction as such, but on the interpersonal causes that might account for it in combination with implications for the future of the relationships. This interdependence-based analysis differs from actor-based approaches, in that they emphasize the actor-and-partner interactions, and what holds them together in the future. For example, sometimes a change (a move) that they initiate and accomplish together may bring about closeness and trust through enhanced interdependence.

Understanding Interpersonal Generosity

Our third example illustrates how *adaptations* might be influenced by *interdependence structure*. That is, the precise properties of interdependence structure are essential to the basic question of under what circumstances generosity might be functional. Our example concerns the best-known and most thoroughly investigated interdependence situation: the prisoner's dilemma. Traditional analyses of situations with this structure have revealed that people enjoy superior outcomes over the course of long-term interaction when they behave on the basis of *quid pro quo*, or tit-for-tat (Axelrod, 1984; Pruitt, 1998): if an interaction partner cooperates, you should likewise cooperate; if a partner competes, you should compete.

But how effective is tit-for-tat under conditions of suboptimal information availability – for example, when people are aware of how a partner's behavior affects their own outcomes, but are not aware of situational constraints that may have shaped the partner's actions? An interdependence analysis suggests that misunderstanding is often rooted in *noise*, or discrepancies between intended outcomes and actual outcomes for a partner that result from unintended errors (Kollock, 1993). For example, when John fails to receive a response to an email message that he sent to Mary, it may be because of a network breakdown in Mary's workplace rather than to Mary's disregard for his well-being. Noise is ubiquitous in everyday interaction, in that the external world is not error-

free (e.g., networks sometimes crash) and people cannot lead error-free lives (e.g., Mary may accidentally delete John's email note in her daily spam purge).

Given that tit-for-tat entails reciprocating a partner's *actual* behavior – and not the partner's *intended* behavior – responding in kind serves to reinforce and exacerbate "accidents." If the accident involves unintended good outcomes, the consequences may be positive. But if the accident entails unintended negative outcomes, the consequences may be more serious. For example, when Mary's actions cause John to suffer poor outcomes, he may respond with tit-for-tat, enacting a behavior that will cause her poor outcomes. In turn – and despite the fact that she did not initially [p. 266 ↓] intend to harm John – Mary will react to John's negative behavior with tit-for-tat, causing him to suffer reciprocal poor outcomes. John and Mary will enter into a pattern of negative reciprocity: they can become trapped in an extended echo effect from which they cannot readily exit – an echo effect that tit-for-tat simply reinforces.

Indeed, research reveals that negative noise exerts detrimental effects when people follow a strict reciprocity rule – partners form more negative impressions of one another and both people suffer poorer outcomes (Van Lange et al., 2002). In contrast, a more generous, tit-for-tat-plus-one strategy (giving the partner a bit more than one received from the partner) yields better outcomes – noise does not negatively affect partners' impressions of one another or the outcomes each receives over the course of interaction. Indeed, in the presence of negative noise, a generous strategy yields better outcomes for both people than does tit-for-tat (for more extended evidence, see Klapwijk and Van Lange, 2009). Such findings are reminiscent of the literature regarding interaction in close relationships, where partners have been shown to enjoy better outcomes in conflictual interactions when one or both partners accommodate or forgive (e.g., Karremans and Van Lange, 2008; Rusbult et al., 1991, see also Simpson, 2007).

The societal implications of this interdependence analysis are quite powerful, as they suggest relatively concrete advice for people entering new situations at school, in organizations, and other situations where people interact in dyads and small groups. Under circumstances of imperfect information (which most situations are like) it helps to give people the benefit of the doubt, to reserve judgment, and to add a little generosity to our tendencies to interact in a tit-for-tat manner. The findings may also be especially relevant to the communication through email, Internet, and other

electronic means, as these devices tend to be quite “noisy.” But perhaps the use of “smileys” and other devices might just serve the very function to communicate trust and generosity to cope with noise.

Understanding Intergroup Relations

Most group phenomena are more complex – often too complex for a comprehensive analysis, which is probably why Thibaut and Kelley often did not go beyond the triad. Nevertheless, the logic provided by interdependence theory has also considerable potential in analyzing intergroup relations.

One important issue is the analysis of intergroup relations. Sometimes groups face high correspondence of outcomes, in that they both (or all) are pursuing the same goal and need each other in that pursuit. For example, neighboring countries help each other in their pursuit of controlling the use of hard drugs. Under such circumstances, groups may actually develop fairly congenial relationships, especially when they hold similar views about the policy for doing so. Sometimes groups face moderate correspondence of outcomes, in the pursuit of some collective goal that is quite costly to each group. For example, countries want to control global warming, but they differ in their interest or views as to how much to contribute. Under such circumstances, groups are faced with social dilemmas (in the intergroup context, a conflict between ingroup interest and common, superordinate interests), and they often exhibit considerably less cooperation than individuals in similar situations (Insko and Schopler, 1998). The primary reasons accounting for that effect are linked to the affordances of the interdependence situation. For example, some degree of conflicting interest challenges trust more (and enhances competitive motivation more) in interactions between groups than between individuals (for a meta-analytic review, see Wildschut et al., 2003). Indeed, there is good deal of evidence that an interdependence approach complements other approaches (such as social identity and self-categorization approaches) in their predictions of intergroup relations.

[p. 267 ↓]

A strong concern with receiving better outcomes – and not getting worse outcomes – than other groups is often conflicting with good outcomes for the collective (De Dreu, 2010). However, competition can sometimes be a powerful means to cooperation. It takes an interdependence approach to analyze the patterns of interdependence between (a) the individual and their group, (b) the individual and the collective, and (c) the group and the collective (see Bornstein, 1992; Halevy et al., 2008; Wit and Kerr, 2002). For example, a soldier (i.e., the individual) who fights forcefully often serves the group (i.e., his/her country), but not necessarily the world (i.e., the entire collective). In such multilayered social dilemmas, competition can be quite beneficial. When there are two (or more) well-defined groups who comprise the entire collective, then sometimes competition between the groups helps the entire collective. The competition should then deal with something desirable. For example, in the Netherlands, there is a contest between cities aiming for the award “Cleanest City.” As another example, two departments at a university may do better (yielding greater research output and enhanced teaching) if the university provides extra resources for only excellent departments. Indeed, organizations often use competition as a means to promote functioning.

Benefits of a Taxonomic Approach: Theoretical Development in the Future

A unique and exceptionally important contribution of interdependence theory is the advancement of a taxonomy of situations. Indeed, there are very few theories in social psychology that advance a taxonomy of situations, even though social psychology as a field is strongly concerned with situational influence or influences from the social environment (see also Reis, 2008). Also, we believe that “dimensions” of temporal structure and information availability that have been added recently (Kelley et al., 2003) will prove to be important to several issues in psychological science and beyond.

First, much research and theory in social psychology focuses on processes in an attempt to understand “system-questions,” such as how cognition and affect might influence each other, the characterization as a dual-process system, such as the reflective and impulsive systems, hot and cool systems, and so on. We suggest that

interdependence theory provides a much-needed taxonomy of situations that may help us understand when (i.e., the situations in which) particular systems might be activated. For example, forms of dependence call for trust, especially when there is some conflict of interest, and perhaps limited time might set into motion a hot system where impulses and gut feelings drive behavior rather than systematic thought (see Hertel et al., 2000). An excellent case in point is the analysis of relationships between “the powerful” and “the powerless” in organizations (Fiske, 1993). Because the latter are strongly dependent on the former, it becomes important to engage in deep, systematical processing for reaching accurate conclusions about the motives and attributes of the powerful. In contrast, the powerful are less dependent on the powerless (and there are often many of the latter), and the powerful are often more shallow, heuristic in forming impressions of the powerless. Accordingly, they are more likely to fall prey to stereotypic information (Fiske, 1993).

Second, a taxonomy of situations is essential to dynamic approaches to social interaction and personality – people do not only respond to situations, they may also actively seek situations, avoid other situations, or shape situations in particular ways (e.g., Snyder and Ickes, 1985). However, it is one thing to recognize that people are not slaves of situational forces – that people select and modify situations in explicit or subtle ways; it is quite another thing to predict the character of situation selection. Interdependence theory provides insight in this respect, in that the dimensions underlying situations should reliably activate and afford specific sorts of [p. 268 ↓] goals and motives. For example, sometimes people may avoid situations of dependence – the decision to work on an independent task rather than a joint task. Situation selection is often functional, in that it helps gratify specific needs or promotes long-term outcomes (Mischel and Shoda, 1995; Snyder and Ickes, 1985). But of course, situation selection may also initiate or sustain self-defeating processes. For example, shy children may avoid interaction, which in turn may limit their opportunities for overcoming shyness. The interdependence theory typology of situations can fruitfully be employed to extend predictive specificity in classic psychological domains, including not only the problem of specificity in predicting how traits relate to situation selection, but also specificity in predicting person-by-situation interactions (Kelley et al., 2003). As such, an interdependence theoretic analysis can advance precise predictions about the inextricable link between persons and situations.

Third, a taxonomic approach is essential to basic evolutionary issues. Because evolutionary theory focuses on the question of how common human characteristics interact with the social environment, it is essential to have the theoretical tools to analyze social situations in terms of their key features (e.g., Schaller et al., 2006; Tooby and Cosmides, 2005; Van Vugt, 2006). Interdependence theory shares some assumptions with evolutionary approaches. One such a shared assumption is that people, as individuals, partners, and as members of a group, *adapt* to social situations (Kelley and Thibaut, 1978). At the same time, while evolutionary theory tends to focus on common human characteristics, interdependence theory can make a contribution by specifying key properties of the social situation to which people adapt – such as the dependence, conflicting interest, information availability, and so on. According to interdependence theory, it is plausible that people develop consistent contingencies, which may take the form of “if … then” rules (Mischel and Shoda, 1999; see Murray and Holmes, 2009; Reis, 2008), in their adaptations to different partners in different social situations. For example, as outlined by Murray and Holmes (2009), if–then rules might reflect the way in which trust is communicated and commitment is built in ongoing relationships – partner's sacrifices might be directly translated into trust. Thus, while evolutionary theory has focused on adaptations, such as coordination and cooperation, interdependence theory provides the conceptual tools for understanding the domains of the situations that afford the expression of the skills and motives relevant to coordination and cooperation. This contribution may be very useful for helping to understand why some cognitions and emotions are closely connected to particular domains of interpersonal situations.

More generally, we suggest that interdependence theory will be exceedingly helpful as a model for understanding when and why particular neurological networks, hormonal responses, or complementary responses might be activated. These biology-based responses will often be adaptive in light of the qualities of both persons and the situation – that is, the SABI model discussed earlier. For example, on the observer's side, responses that are linked to anger are probably best understood when carefully analyzing another person's violation of a norm in situations where people are likely to have somewhat conflicting preferences (e.g., Singer et al., 2006). It is especially striking that people with prosocial orientations tend to react very automatically to a violation of equality (e.g., activation in the amygdala, Haruno and Frith, 2009). Such findings

provide neuroscientific evidence in support of the integrative model of social value orientation, which states that prosocial orientation represents not only the tendency to enhance joint outcomes but also the tendency to enhance equality in outcomes (Van Lange, 1999). On the actor's side, feelings of guilt might be evoked in such situations when we ourselves violate such norms (e.g., Pinter et al., 2007). Further, the topic of self-regulation (and affect-regulation and self-control) in the interpersonal domain is of course strongly linked to inhibiting the [p. 269 ↓] temptation of self-interest and exercising self-restraint.

Conclusion

Social psychology is the field of psychology that is defined most strongly in terms of influences of the situation – specifically, the influence of the social environment on human behavior. Somewhat surprisingly, not much theorizing in social psychology is centered on the analysis of the social environment. By providing a taxonomy of interpersonal situations, interdependence theory has served that role. The addition of new dimensions (information availability and temporal structure) to the well-established ones (dependence, mutuality of dependence, basis of dependence, covariation of interest) should be essential toward understanding the nature and mechanics of (implicit) theories that people bring to bear on situations with limited information (e.g., the hot and cold systems, the degree of processing, the needs and motives involved, as well as the implicit theories by which people make incomplete information complete) as well as the motives and skills that are relevant to time in a general sense (e.g., investment, delay of gratification, consideration for future consequences). A taxonomy of interpersonal situation is essential for theoretical progress.

From a theoretical perspective, it is crucial that we need to know better what a situation "objectively" represents, because only then it is possible to understand what people subjectively make of a situation (construction). Conceptually, the constructs of given situation (objective situation), transformation (meaning analysis), and effective situation (subjective situation) represent the heart of the interdependence theory. It complements much other theorizing in social psychology which tends to focus on the processes relevant to transformation and effective situation preferences. Another reason why a taxonomy is important is that it helps us understand the situations that people might

face (in terms of valence, frequency, and intensity), and how these features covary with several factors, such as differences in personality, social class, gender, and age. For example, the frequency with which one faces situations of unilateral dependence on another person might increase from adulthood to old age. A taxonomy of situations also helps us understand the situations that relationship partners and members of small groups are likely to face (or not) – for example, of how they face situation of conflicting interests. As a variation of Lewin's (1952: 169) well-known dictum, one might suggest that "there is nothing as practical as a good taxonomy."

Thus, after more than 50 years since Thibaut and Kelley (1959), interdependence theory comes full circle. It really has helped the field to get a grip on situations that interacting partners face or might face (the given interdependence situation), what they make of it (the transformation process) in terms of cognition and emotion, and how the structure and the processes shape human behavior and social interactions. This also helps to explain why interdependence theory has been well appreciated for over five decades, and why interdependence has been used to understand so many issues – group dynamics, power and dependence, social comparison, conflict and cooperation, attribution and self-presentation, trust and distrust, emotions, love and commitment, coordination and communication, risk and self-regulation, performance and motivation, social development, and neuroscientific models of social interaction. We are looking forward to the theoretical contributions and applications of interdependence theory over the next 50 years.

Note

1 Sadly, Caryl Rusbult (1952–2010) passed away on January 27, 2010, some weeks before this chapter was completed. Some of her important contributions to interdependence theory throughout her career are described in this chapter. We thank John [p. 270 ↓] Holmes, Norbert Kerr and Harry Reis for their helpful comments on an earlier draft of this chapter.

Paul A.M.Van Lange Caryl E.Rusbult

References

- Agnew, C.R., Van Lange, P.A.M., Rusbult, C.E. and Langston, C.A. Cognitive interdependence: Commitment and the mental representation of close relationships . Journal of Personality and Social Psychology , 74, 939–954. (1998) <http://dx.doi.org/10.1037/0022-3514.74.4.939>
- Aguilar, R.J. and Nightingale, N.N. The impact of specific battering experiences on the self-esteem of abused women . Journal of Family Violence , 9, 35–45. (1994) <http://dx.doi.org/10.1007/BF01531967>
- Ajzen, I. The theory of planned behavior . Organizational Behavior and Human Decision Processes , 50, 179–211. (1991) <http://dx.doi.org/10.1016/0749-5978%2891%2990020-T>
- Andersen, S.M. and Chen, S. The relational self: An interpersonal social-cognitive theory . Psychological Review , 109, 619–645. (2002) <http://dx.doi.org/10.1037/0033-295X.109.4.619>
- Attridge, M., Berscheid, E. and Simpson, J.A. Predicting relationship stability from both partners versus one . Journal of Personality and Social Psychology , 69, 254–268. (1995) <http://dx.doi.org/10.1037/0022-3514.69.2.254>
- Axelrod, R. (1984) The Evolution of Cooperation . New York: Basic Books.
- Baron, R.S. and Kerr, N.L. (2003) Group Process, Group Decision, Group Action , 2nd Edition. Buckingham: Open University Press.
- Baumeister, R.F. and Leary, M.R. The need to belong: Desire for interpersonal attachments as a fundamental human motivation . Psychological Bulletin , 117, 497–529. (1995) <http://dx.doi.org/10.1037/0033-2909.117.3.497>
- Bornstein, G. The free rider problem in intergroup conflicts over step-level and continuous public goods . Journal of Personality and Social Psychology , 62, 597–606. (1992) <http://dx.doi.org/10.1037/0022-3514.62.4.597>

Carver, C.S. and Scheier, M.F. (1998) On the Self-regulation of Behavior . New York: Cambridge.

Clark, M.S., Dubash, P. and Mills, J. Interest in another's consideration of one's needs in communal and exchange relationships . *Journal of Experimental Social Psychology* , 34, 246–264. (1998) <http://dx.doi.org/10.1006/jesp.1998.1352>

Collins, N.L. and Feeney, B.C. Working models of attachment shape perceptions of social support: Evidence from experimental and observational studies . *Journal of Personality and Social Psychology* , 87, 363–383. (2004) <http://dx.doi.org/10.1037/0022-3514.87.3.363>

Collins, N.L. and Miller, L.C. Self-disclosure and liking: A meta-analytic review . *Psychological Bulletin* , 116, 457–475. (1994) <http://dx.doi.org/10.1037/0033-2909.116.3.457>

De Dreu, C.K.W. (2010) Social conflict: The emergence and consequences of struggle and negotiation . In S.T. Fiske, ed. , D.T. Gilbert, ed. and G. Lindzey (eds), *Handbook of Social Psychology*, Vol. 2 , 5th Edition, pp. 983–1023. New York: Wiley.

Deutsch, M. Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social Issues* , 31, 137–149. (1975) <http://dx.doi.org/10.1111/j.1540-4560.1975.tb01000.x>

Finkel, E.J., Campbell, W.K., Brannel, A.B., Dalton, A.N., Scarbeck, S.J. and Chartrand, T.L. High-maintenance interaction: Inefficient social coordination impairs self-regulation . *Journal of Personality and Social Psychology* , 91, 456–475. (2006) <http://dx.doi.org/10.1037/0022-3514.91.3.456>

Finkel, E.J., Rusbult, C.E., Kumashiro, M. and Hannon, P.A. Dealing with betrayal in close relationships: Does commitment promote forgiveness? *Journal of Personality and Social Psychology* , 82, 956–974. (2002) <http://dx.doi.org/10.1037/0022-3514.82.6.956>

Fiske, A.P. The four elementary forms of sociality: Framework for a unified theory of social relations . *Psychological Review* , 99, 689–723. (1992) <http://dx.doi.org/10.1037/0033-295X.99.4.689>

Fiske, S.T. Controlling other people: The impact of power on stereotyping . American Psychologist , 48, 621–628. (1993) <http://dx.doi.org/10.1037/0003-066X.48.6.621>

Fiske, S.T. (2004). Social Beings: A Core Motives Approach to Social Psychology . New York: Wiley.

Fraley, R.C. and Shaver, P.R. Adult romantic attachment: Theoretical developments, emerging controversies, and unanswered questions . Review of General Psychology , 4, 132–154. (2000) <http://dx.doi.org/10.1037/1089-2680.4.2.132>

Greenwald, A.G., McGhee, D.E. and Schwartz, J.L.K. Measuring individual differences in implicit cognition: The implicit association test . Journal of Personality and Social Psychology , 74, 1464–1480. (1998) <http://dx.doi.org/10.1037/0022-3514.74.6.1464>

Halevy, N., Bornstein, G. and Sagiv, L. ‘Ingroup love’ and ‘Outgroup hate’ as motives for individual participation in intergroup conflict: A new game paradigm . Psychological Science , 19, 405–411. (2008) <http://dx.doi.org/10.2307/2695887>

Haruno, M. and Frith, C.D. Activity in the amygdala elicited by unfair divisions predicts social value orientation . Nature Neuroscience , 13, 160–161. (2009) <http://dx.doi.org/10.1038/nn.2468>

Hertel, G., Neuhof, J., Theuer, T. and Kerr, N. Mood effects on cooperation in small groups: Does positive mood simply lead to more cooperation? Cognition and Emotion , 14, 441–472. (2000) <http://dx.doi.org/10.1080/026999300402754>

Higgins, E.T. Beyond pleasure and pain . American Psychologist , 52, 1280–1300. (1997) <http://dx.doi.org/10.1037/0003-066X.52.12.1280>

Higgins, E.T. Making a good decision: Value from fit . American Psychologist , 55, 1217–1230. (2000) <http://dx.doi.org/10.1037/0003-066X.55.11.1217>

Higgins, E.T. (2011) Regulatory focus theory . In P.A.M. Van Lange, ed. , A.W. Kruglanksi, ed. and E.T. Higgins (eds), Handbook of Theories of Social Psychology, Vol. 1 . London: Sage <http://dx.doi.org/10.4135/9781446249215>

Holmes, J.G. Social relationships: The nature and function of relational schemas . European Journal of Social Psychology , 30, 447–495. (2002)
<http://dx.doi.org/10.1002/1099-0992%28200007/08%2930:4%3C447::AID-EJSP10%3E3.0.CO;2-Q>

Holmes, J.G. The benefits of abstract functional analysis in theory construction: The case of interdependence theory . Personality and Social Psychology Review , 8, 146–155. (2004) http://dx.doi.org/10.1207/s15327957pspr0802_8

Holmes, J.G. and Murray, S.L. (1996) Conflict in close relationships . In E.T. Higgins, ed. and A. Kruglanski (eds), Social Psychology: Handbook of Basic Principles , pp. 622–654. New York: Guilford Press.

Holmes, J.G. and Rempel, J.K. (1989) Trust in close relationships . In C. Hendrick (ed.), Review of Personality and Social Psychology , 10, 187–220. London: Sage.

Homans, G.C. (1950) The Human Group . New York: Harcourt, Brace & World.

Insko, C.A. and Schopler, J. (1998) Differential distrust of groups and individuals . In C. Sedikides, ed. , J. Schopler, ed. and C.A. Insko (eds), Intergroup Cognition and Intergroup Behavior: Toward a Closer Union , pp. 75–107. Hillsdale, NJ: Erlbaum.

Jones, E.J. (1998) Major developments in five decades of social psychology . In D. Gilbert, ed. , S. Fiske, ed. and G. Lindzey (eds), Handbook of Social Psychology, Vol. 2 , 4th Edition, pp. 3–57. Boston: McGraw-Hill.

Karremans, J.C. and Van Lange, P.A.M. Forgiveness in personal relationships: Its malleability and powerful consequences . European Review of Social Psychology , 19, 202–241. (2008) <http://dx.doi.org/10.1080/10463280802402609>

Kelley, H.H. The situational origins of human tendencies: A further reason for the formal analysis of structures . Personality and Social Psychology Bulletin , 9, 8–30. (1983)
<http://dx.doi.org/10.1177/0146167283091003>

Kelley, H.H. The theoretical description of interdependence by means of transition lists . *Journal of Personality and Social Psychology* , 47, 956–982. (1984) <http://dx.doi.org/10.1037/0022-3514.47.5.956>

Kelley, H.H., Holmes, J.G., Kerr, N.L., Reis, H.T., Rusbult, C.E. and Van Lange, P.A.M. (2003) *An Atlas of Interpersonal Situations* . New York: Cambridge.

Kelley, H.H. and Stahelski, A.J. Social interaction basis of cooperators' and competitors' beliefs about others . *Journal of Personality and Social Psychology* , 16, 66–91. (1970) <http://dx.doi.org/10.1037/h0029849>

Kelley, H.H. and Thibaut, J.W. (1978) *Interpersonal Relations: A Theory of Interdependence* . New York: Wiley.

Kerr, N.L. and Tindale, R.S. Small group decision making and performance . *Annual Review of Psychology* , 55, 623–656. (2004) <http://dx.doi.org/10.1146/annurev.psych.55.090902.142009>

Klapwijk, A. and Van Lange, P.A.M. Promoting cooperation and trust in 'noisy' situations: The power of generosity . *Journal of Personality and Social Psychology* , 96, 83–103. (2009) <http://dx.doi.org/10.1037/a0012823>

Kollock, P. 'An eye for an eye leaves everyone blind': Cooperation and accounting systems . *American Sociological Review* , 58, 768–786. (1993) <http://dx.doi.org/10.2307/2095950>

Lewin, K. (1952) *Field Theory in Social Sciences: Selected Theoretical Papers* . New York: Harper.

Luce, R.D. and Raiffa, H. (1957) *Games and Decisions: Introduction and Critical Survey* . London: Wiley.

Messick, D.M. and McClintock, C.G. Motivational bases of choice in experimental games . *Journal of Experimental Social Psychology* , 4, 1–25. (1968) <http://dx.doi.org/10.1016/0022-1031%2868%2990046-2>

Miller, R.S. Inattentive and contented: Relationship commitment and attention to alternatives . *Journal of Personality and Social Psychology* , 73, 758–766. (1997) <http://dx.doi.org/10.1037/0022-3514.73.4.758>

Mischel, W. and Shoda, Y. A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, and invariance in personality structure . *Psychological Review* , 102, 246–268. (1995) <http://dx.doi.org/10.1037/0033-295X.102.2.246>

Murray, S.L. and Holmes, J.G. The architecture of interdependent minds: A Motivation-management theory of mutual responsiveness . *Psychological Review* , 116, 908–928. (2009) <http://dx.doi.org/10.1037/a0017015>

Murray, S.L., Holmes, J.G. and Collins, N.L. Optimizing assurance: The risk regulation system in relationships . *Psychological Bulletin* , 132, 641–666. (2006) <http://dx.doi.org/10.1037/0033-2909.132.5.641>

Nisbett, R.E. and Cohen, D. (1996) Culture of Honor: The Psychology of Violence in the South . Boulder, CO: Westview.

Pinter, B., Insko, C.A., Wildschut, T., Kirchner, J.L., Montoya, R.M. and Wolf, S.T. Reduction of the interindividual-intergroup discontinuity: The role of leader accountability and proneness to guilt . *Journal of Personality and Social Psychology* , 93, 250–265. (2007) <http://dx.doi.org/10.1037/0022-3514.93.2.250>

Pruitt, D. (1998) Social conflict . In D. Gilbert, ed. , S. Fiske, ed. and G. Lindzey (eds), *Handbook of Social Psychology*, Vol. 2 , 4th Edition, pp. 470–503. Boston: McGraw-Hill.

Reis, H.T. Reinvigorating the concept of situation in social psychology . *Personality and Social Psychology Review* , 12, 311–329. (2008) <http://dx.doi.org/10.1177/1088868308321721>

Reis, H.T., Collins, W.A. and Berscheid, E. The relationship context of human behavior and development . *Psychological Bulletin* , 126, 844–872. (2000) <http://dx.doi.org/10.1037/0033-2909.126.6.844>

Righetti, F. and Rusbult, C.E. (2007) Interpersonal regulatory fit: Consequences for goal pursuit . Unpublished manuscript, Vrije Universiteit Amsterdam.

Rusbult, C.E. A longitudinal test of the investment model: The development (and deterioration) of satisfaction and commitment in heterosexual involvements . Journal of Personality and Social Psychology , 45, 101–117. (1983) <http://dx.doi.org/10.1037/0022-3514.45.1.101>

Rusbult, C.E., Coolsen, M.K., Kirchner, J.L. and Clarke, J. (2006) Commitment . In A. Vangelisti, ed. and D. Perlman (eds), Handbook of Personal Relationships , pp. 615–635. New York: Cambridge University Press.

Rusbult, C.E. and Martz, J.M. Remaining in an abusive relationship: An investment model analysis of nonvoluntary commitment . Personality and Social Psychology Bulletin , 21, 558–571. (1995) <http://dx.doi.org/10.1177/0146167295216002>

Rusbult, C.E. and Van Lange, P.A.M. (1996) Interdependence processes . In E.T. Higgins, ed. and A. Kruglanski (eds), Social Psychology: Handbook of Basic Principles , pp. 564–596. New York: Guilford Press.

Rusbult, C.E. and Van Lange, P.A.M. Interdependence, interaction, and relationships . Annual Review of Psychology , 54, 351–375. (2003) <http://dx.doi.org/10.1146/annurev.psych.54.101601.145059>

Rusbult, C.E., Verette, J., Whitney, G.A., Slovik, L.F. and Lipkus, I. Accommodation processes in close relationships: Theory and preliminary empirical evidence . Journal of Personality and Social Psychology , 60, 53–78. (1991) <http://dx.doi.org/10.1037/0022-3514.60.1.53>

Schaller, M., ed. , Kenrick, D., ed. and Simpson J. (eds) (2006) Evolution and Social Psychology . New York: Psychology Press.

Simpson, J.A. Psychological foundations of trust . Current Directions in Psychological Science , 16, 264–268. (2007) <http://dx.doi.org/10.1111/j.1467-8721.2007.00517.x>

Singer, T., SeymourB., O'DohertyJ., KlaasE.S., DolanJ.D. and Frith, C. Empathic neural responses are modulated by the perceived fairness of others . *Nature* , 439, 466–469. (2006) <http://dx.doi.org/10.1038/nature04271>

Snyder, M. and Ickes, W. (1985) Personality and social behavior . In G. Lindzey, ed. and E. Aronson (eds), *The Handbook of Social Psychology* , pp. 883–947. New York: Random House.

Surra, C.A. and Longstreth, M. Similarity of outcomes, interdependence, and conflict in dating relationships . *Journal of Personality and Social Psychology* , 59, 501–516. (1990) <http://dx.doi.org/10.1037/0022-3514.59.3.501>

Thibaut, J.W. and Kelley, H.H. (1959) *The Social Psychology of Groups* . New York: Wiley.

Tooby, J. and Cosmides, L. (2005) Conceptual foundations of evolutionary psychology . In D.M. Buss (ed.), *The Handbook of Evolutionary Psychology* , pp. 5–67. Hoboken, NJ: Wiley.

Turiel, E. (1983) *The Development of Social Knowledge: Morality and Convention* . Cambridge: Cambridge University Press.

Van Lange, P.A.M. The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation . *Journal of Personality and Social Psychology* , 77, 337–349. (1999) <http://dx.doi.org/10.1037/0022-3514.77.2.337>

Van Lange, P.A.M., De Cremer, D., Van Dijk, E. and Van Vugt, M. (2007) Self-interest and beyond: Basic principles of social interaction . In A.W. Kruglanski, ed. and E.T. Higgins (eds), *Social Psychology: Handbook of Basic Principles* , pp. 540–561. New York: Guilford Press.

Van Lange, P.A.M. and Joireman, J.A. How can we promote behaviour that serves all of us in the future . *Social Issue and Policy Review* , 2, 127–157. (2008) <http://dx.doi.org/10.1111/j.1751-2409.2008.00013.x>

Van Lange, P.A.M., Otten, W., De Bruin, E.M.N. and Joireman, J.A. Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence . *Journal of Personality and Social Psychology* , 73, 733–746. (1997) <http://dx.doi.org/10.1037/0022-3514.73.4.733>

Van Lange, P.A.M., Ouwerkerk, J.W. and Tazelaar, M.J.A. How to overcome the detrimental effects of noise in social interaction: The benefits of generosity . *Journal of Personality and Social Psychology* , 82, 768–780. (2002) <http://dx.doi.org/10.1037/0022-3514.82.5.768>

Van Lange, P.A.M., Rusbult, C.E., Drigotas, S.M., Arriaga, X.B., Witcher, B.S. and Cox, C.L. Willingness to sacrifice in close relationships . *Journal of Personality and Social Psychology* , 72, 1373–1395. (1997) <http://dx.doi.org/10.1037/0022-3514.72.6.1373>

Van Vugt, M. Evolutionary origins of leadership and followership . *Personality and Social Psychology Review* , 10, 354–372. (2006) http://dx.doi.org/10.1207/s15327957pspr1004_5

Vuolevi, J.H.K. and Van Lange, P.A.M. Beyond the information given: The power of the belief in self-interest . *European Journal of Social Psychology* , 40, 26–34. (2010)

Walker, L. (2000) *The Battered Woman Syndrome* , 2nd Edition. New York: Springer.

Wildschut, T., Pinter, B., Vevea, J.L., Insko, C.A. and Schopler, J. Beyond the group mind: A quantitative review of the interindividual-intergroup discontinuity effect . *Psychological Bulletin* , 129, 698–722. (2003) <http://dx.doi.org/10.1037/0033-2909.129.5.698>

Wit, A.P. and Kerr, N.L. 'Me vs. just us vs. us all' Categorization and cooperation in nested social dilemmas . *Journal of Personality and Social Psychology* , 83, 616–637. (2002) <http://dx.doi.org/10.1037/0022-3514.83.3.616>

<http://dx.doi.org/10.4135/9781446249222.n39>

Handbook of Theories of Social Psychology

A Theory of Cooperation – Competition and beyond

Contributors: Morton Deutsch

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "A Theory of Cooperation – Competition and beyond"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n40>

Print pages: 275-295

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n40>

[p. 275 ↓]

Chapter 40: A Theory of Cooperation – Competition and beyond

Morton Deutsch

Abstract

This chapter is concerned with my inter-related theoretical work in the areas of cooperation – competition, conflict resolution, social justice, and social relations. The theory of *cooperation–competition* is a component of the other theories. Thus, the theory of *conflict resolution* is based on this theory and my Crude Law of Social Relations. My work in *social justice* is also based on this theory, the Crude Law, and on my theoretical work dealing with social relations. The work in *social relations* sketches a more generalized approach to the understanding of the bidirectional interaction between social relations and psychological orientations. In a prelude to my theoretical discussion, I consider the personal, social, and professional influences that shaped my work. In the last part, I describe some of the important social effects of this work.

Introduction

In this chapter, I shall present my interrelated theoretical work on cooperation – competition, conflict resolution, social justice, and social relations. I shall omit a presentation of relevant research since this has been presented elsewhere. Thus, Johnson and Johnson (2005) in their excellent monograph, *New Developments in Social Interdependence Theory*, present an extensive summary of relevant research and social practice that relate to this theory. In books related to conflict resolution (Deutsch, 1973; Deutsch et al., 2006) there is considerable discussion of research.

Similarly, in Deutsch (1985) and *Social Justice Research* (Vol. 19[1], March 2006) there is presentation of research related to social justice.

This chapter is divided into three parts. The first is concerned with the personal, social, and theoretical influences affecting the development of my theoretical and research work; the second presents my basic theoretical ideas; and the third discusses the social effects of my work.

Influences on My Work: Personal, Social, and Theoretical

The choice of areas for social psychological work is affected not only by professional and [p. 276 ↓] scientific contacts and readings, but also by personal and familial experiences as well as by broader social and cultural influences. In my case, being Jewish and the youngest child in my family, in my school, and in my neighborhood group exposed me to considerable prejudice¹ and put-downs. These experiences sensitized me to prejudice and led to an identification with underdogs.

At age 15, I entered the City College of New York (CCNY) in 1935 as a premed major with the idea of becoming a psychiatrist, having been intrigued by the writings of Sigmund Freud, some of which I had read before college. I was drawn to psychoanalysis undoubtedly because it appeared to be so relevant to personal issues with which I was struggling, and also because it was so radical (it seemed to be so in the early and mid 1930s). During my adolescence, I was also politically radical and somewhat rebellious toward authority, helping to organize a student strike against terrible food in the Townsend Harris High School lunchroom, and later, a strike against the summer resort owners who were exploiting the college student waiters and bus boys at Camp Copake, of whom I was one.

The 1930s were a turbulent period, internationally as well as domestically. The economic depression; labor unrest; the rise of Nazism and other forms of totalitarianism; the Spanish civil war; the ideas of Marx, Freud, and Einstein; as well as the impending Second World War were shaping the intellectual atmosphere that affected psychology.

Several members of the psychology faculty at CCNY were active in creating the Psychologist League, the precursor to the Society for the Psychological Study of Social Issues. Thus, when I became disenchanted with the idea of being a premed student (after dissecting a pig in a biology lab), I was happy to switch to a psychology major: it was a simpatico faculty. Psychology was a part of the Department of Philosophy at CCNY when I started my major in it. Morris Raphael Cohen, the distinguished philosopher of science, was the leading intellectual figure at CCNY, and his influence permeated the atmosphere. I note that in the lunchroom alcoves at CCNY, I became well versed in Marxist theology and disputation. Students adhering to the First, Second, Third, and Fourth International congregated in different corners of the lunchroom.

My first exposure to Lewin's writings was in two undergraduate courses, taken simultaneously: social psychology, and personality and motivation. In these courses, I read Lewin's *Dynamic Theory of Personality* (1935), *Principles of Topological Psychology* (1936), and "The conceptual representation and measurement of psychological forces" (1938). I and others experienced great intellectual excitement on reading these books more than 60 years ago. These books are permeated by a view of the nature of psychological science different from the traditional. The new view was characterized by Lewin as the *Galilean mode of thought*, which was contrasted with the classic *Aristotelian mode*. In my writings on field theory (Deutsch, 1954), I have characterized in some detail Lewin's approach to psychological theorizing – his metatheory.

Although I was impressed by Lewin's writing, my career aspirations were still focused on becoming a psychoanalytic psychologist as I decided to do graduate work. My undergraduate experiences, in as well as outside the classroom, led me to believe that an integration of psychoanalysis, Marxism, and scientific method, as exemplified by Lewin's work, could be achieved. In the 1930s such influential figures as Wilhelm Reich, Erich Fromm, Max Horkheimer, Theodor Adorno, and Else Frenkel-Brunswik were trying to develop an integration of psychoanalysis and Marxism. Also at this time, some psychoanalytic theorists such as David Rappaport were intrigued by the idea that the research conducted by Lewin and his students on tension systems could be viewed as a form of experimental psychoanalysis.

After obtaining my MA in 1940 in clinical psychology at the University of Pennsylvania, I started a rotating clinical internship at three [p. 277 ↓] New York State institutions: one was for the feeble-minded (Letchworth Village), another for delinquent boys (Warwick), and a third for psychotic children as well as adults (Rockland State Hospital). When the attack on Pearl Harbor occurred on December 7, 1941, I was still in my psychology internship. Shortly thereafter, I joined the Air Force. I flew in 30 bombing missions against Germany. During combat, I saw many of our planes as well as German planes shot down, and I also saw massive damage inflicted by our bombs and those of the Royal Air Force on occupied Europe and Germany. Moreover, being stationed in England, I saw the great destruction wreaked by the German air raids and felt common apprehensions while sitting in air-raid shelters during German bombings. Although I had no doubt of the justness of the war against the Nazis, I was appalled by its destructiveness.

After my demobilization, I decided to apply for admission to the doctoral programs at the University of Chicago (where Carl Rogers and L.L. Thurstone were the leading lights), at Yale University (where Donald Marquis was chairman and where Clark Hull was the major attraction), and at MIT (where Kurt Lewin had established a new graduate program and the Research Center for Group Dynamics). As one of the first of the returning soldiers, I had no trouble getting interviews or admissions to all three schools. I was most impressed by Kurt Lewin and his vision of his newly established Research Center and so decided to take my PhD at MIT. I date the start of my career as a social psychologist to my first meeting with Lewin, in which I was enthralled by him and committed myself to studying at his Research Center for Group Dynamics (RCGD).

Lewin assembled a remarkable group of faculty and students to compose the RCGD at MIT. For the faculty, he initially recruited Dorwin Cartwright, Leon Festinger, Ronald Lippitt, and Marian Radke (later Radke-Yarrow). Jack French and Alvin Zander were to join subsequently. The small group of students included Kurt Back, Alex Bavelas, David Emery, Gordon Hearn, Murray Horowitz, David Jenkins, Harold Kelley, Albert Pepitone, Stanley Schachter, Richard Snyder, John Thibaut, Ben Willerman, and myself. These faculty members and students were extraordinarily productive, and they played a pivotal role in developing modern social psychology in its applied as well as its basic aspects.

My career in social psychology has been greatly affected by Kurt Lewin and my experiences at the RCGD. The intellectual atmosphere created by Lewin strongly shaped my dissertation and my value orientation as a social psychologist. Lewin was not only an original, tough-minded theorist and researcher with a profound interest in the philosophy and methodology of science, he was also a tender-hearted psychologist who was deeply involved with developing psychological knowledge that would be relevant to important human concerns (Deutsch, 1992). He provided a scientific role model that I have tried to emulate. Like Lewin, I have wanted my theory and research to be relevant to important social issues, but I also wanted my work to be scientifically rigorous and tough-minded. As a student, I was drawn to both the tough-mindedness of Festinger's work and to the direct social relevance of Lippitt's approach and did not feel the need to identify with one and derogate the other.

My dissertation started off with an interest in issues of war and peace (atomic bombs had been dropped on Hiroshima and Nagasaki shortly before I resumed my graduate studies) and with an image of the possible ways that the nations composing the newly formed United Nations Security Council would interact. The atmosphere at the Center, still persisting after Lewin's premature death, led me to turn this social concern about the risk of nuclear war into a theoretically oriented, experimental investigation of the effects of cooperative and competitive processes. The specific problem that I was first interested in took on a more generalized form. It had been transformed into an attempt to understand the fundamental features of cooperative and [p. 278 ↓] competitive relations and the consequences of these different types of interdependencies in a way that would be generally applicable to the relations among individuals, groups, or nations. The problem had become a theoretical one, with the broad scientific goal of attempting to develop insight into a variety of phenomena through several fundamental concepts and basic propositions. The intellectual atmosphere at the Center pushed its students to theory building.

As I reflect back on the intellectual roots of my dissertation, I see it was influenced not only by Lewin's theoretical interest in social interdependence but also by Marxist concerns with two different systems of distributive justice: a cooperative-egalitarian and a competitive-meritocratic one. In addition, the writing of George Herbert Mead (1934) affected my way of thinking about cooperation and its importance to human development. Also, my discussion of the relation between objective social

interdependence and perceived social interdependence was much influenced by Koffka's (1935) answer to the question "Why do things look as they do?" Further, my reading of the existing literature on cooperation – competition (Barnard, 1938; Lewin, 1948; Maller, 1929; May and Doob, 1937; G.H. Mead, 1934; M. Mead, 1937) indirectly influenced my work.

Theory

This part is concerned with my inter-related theoretical work in the areas of cooperation – competition, conflict resolution, social justice, and social relations. The theory of *cooperation–competition* is a component of the other theories. Thus, the theory of *conflict resolution* is based on this theory and my Crude Law of Social Relations. My work in *social justice* is also based on this theory, the Crude Law, and on my theoretical work dealing with social relations. The work in *social relations* sketches a more generalized approach to the understanding of the bidirectional interaction between social relations and psychological orientations. Social relations are characterized not only by the dimension of cooperation–competition but also by such other dimensions as: equality–inequality; task-oriented versus social–emotional-oriented (*Gemeinschaft* versus *Gesellschaft*); formal versus informal; and degree of importance. Psychological orientations include the following components: cognitive, motivational, moral, and action orientations.

Before discussing this work, let me note that I recognize that my *theories* in social psychology have considerable ambiguity inherent in them and lack the precision of the theories in the natural sciences. For example, in my theoretical and experimental work on cooperation–competition, I discussed the problem of the relation between the "objective" and "perceived" reality of social interdependence; a similar problem confronts all experimental social psychologists.

A Theory of Cooperation and Competition

In my 1949 presentation of my theory of cooperation and competition, I employed the Lewinian terminology related to locomotion in the life space and developed the

hypotheses of the theory in a formal, hypothetico-deductive manner. I think the ideas were fine, but the presentation was awkward and the language too idiosyncratic. In more recent presentations, including this one, I have presented the theory in a more accessible, informal manner.

The theory has two basic ideas. One relates to the type of interdependence among goals of the people involved in a given situation. The other pertains to the type of action taken by the people involved. I identify two basic types of goal interdependence: positive (where the goals are linked in such a way that the amount or probability of a person's goal attainment is positively correlated with the amount or probability of another obtaining his goal) and negative (where the goals [p. 279 ↓] are linked in such a way that the amount or probability of goal attainment is negatively correlated with the amount or probability of the other's goal attainment). To put it colloquially, if you're positively linked with another, then you sink and swim together; with negative linkage, if the other sinks, you swim, and if the other swims, you sink.

It is well to realize that few situations are "purely" positive or negative. In most situations, people have a mixture of goals so that it is common for some of their goals initially to be positive and some negatively interdependent. In this section, for analytical purposes, I discuss pure situations. In conflict and other mixed situations, the relative strength of the two types of goal interdependence, as well as the parties' general orientation to one another, largely determine the nature of their interaction.

I also characterize two basic types of action by an individual: "effective actions," which improve the actor's chances of obtaining a goal, and "bungling actions," which worsen the actor's chances of obtaining the goal. (For the purpose of simplicity, I use dichotomies for my basic concepts; the dichotomous types of interdependence and the dichotomous types of actions are, I assume, polar ends of continua.) I then combine types of interdependence and types of action to posit how they jointly affect three basic social psychological processes that are discussed later: *substitutability*, *cathexis*, and *inducibility*.

People's goals may be linked for various reasons. Thus, positive interdependence can result from people liking one another; being rewarded in terms of their joint achievement, needing to share a resource or overcome an obstacle together, holding

common membership or identification with a group whose fate is important to them, being unable to achieve their task goals unless they divide up the work, being influenced by personality and cultural orientation, being bound together because they are treated this way by a common enemy or an authority, and so on. Similarly, with regard to negative interdependence, it can result from people disliking one another or from their being rewarded in such a way that the more the other gets of the reward, the less one gets, and so on.

In addition to positive and negative interdependence, it is well to recognize that there can be lack of interdependence, or *independence*, such that the activities and fate of the people involved do not affect one another, directly or indirectly. If they are completely independent of one another, no conflict arises; the existence of a conflict implies some form of interdependence.

One further point: asymmetries may exist with regard to the degree of interdependence in a relationship; suppose that what you do or what happens to you may have a considerable effect on me, but what I do or what happens to me may have little impact on you. I am more dependent on you than you are on me. In the extreme case, you may be completely independent of me and I may be highly dependent on you. As a consequence of this asymmetry, you have greater power and influence in the relationship than I. This power may be general if the asymmetry exists in many situations, or it may be situation-specific if the asymmetry occurs only in a particular situation. A master has general power over a slave, while an auto mechanic repairing my car's electrical system has situation-specific power.

The three concepts mentioned previously – *substitutability*, *cathexis*, and *inducibility* – are vital to understanding the social psychological processes involved in creating the major effects of cooperation and competition. *Substitutability* (how a person's actions can satisfy another person's intentions) is central to the functioning of all social institutions (the family, industry, schools), to the division of labor, and to role specialization. Unless the activities of other people can substitute for yours, you are like a person stranded on a desert island alone: you have to build your own house, find or produce your own food, protect yourself from harmful animals, treat your ailments and illnesses, educate yourself about the nature of your new [p. 280 ↓] environment and about how to do all these tasks, and so on, without the help of others. Being alone,

you can neither create children nor have a family. *Substitutability* permits you to accept the activities of others in fulfilling your needs. *Negative substitutability* involves active rejection and effort to counteract the effects of another's activities.

Cathexis refers to the predisposition to respond evaluatively, favorably, or unfavorably to aspects of one's environment or self. Through natural selection, evolution has ensured that all living creatures have the capacity to respond positively to stimuli that are beneficial to them and negatively to those that are harmful. They are attracted to, approach, receive, ingest, like, enhance, and otherwise act positively toward beneficial objects, events, or other creatures; in contrast, they are repelled by harmful objects and circumstances and avoid, eject, attack, dislike, negate, and otherwise act negatively toward them. This inborn tendency to act positively toward the beneficial and negatively toward the harmful is the foundation on which the human potentials for cooperation and love as well as for competition and hate develop. The basic psychological orientation of cooperation implies the positive attitude, "We are for each other," "We benefit one another;" competition, by contrast, implies the negative attitude "We are against one another" and, in its extreme form, "You are out to harm me."

Inducibility refers to the readiness to accept another's influence to do what he or she wants; negative inducibility refers to the readiness to reject or obstruct fulfillment of what the other wants. The complement of *substitutability* is *inducibility*. You are willing to be helpful to another whose actions are helpful to you, but not to someone whose actions are harmful. In fact, you reject any request to help the other engage in harmful actions and, if possible, obstruct or interfere with these actions.

The Effects of Cooperation and Competition

Thus, the theory predicts that if you are in a positive interdependent relationship with someone who bungles, the bungling is not a substitute for effective actions you intended; thus, the bungling is viewed negatively. In fact, when your net-playing tennis partner in a doubles game allows an easy shot to get past him, you have to extend yourself to prevent being harmed by the error. On the other hand, if your relationship is one of negative interdependence, and the other person bungles (as when your tennis opponent double-faults), your opponent's bungle substitutes for an effective

action on your part, and it is regarded positively or valued. The reverse is true for effective actions. An opponent's effective actions are not substitutable for yours and are negatively valued; a teammate can induce you to help him or her make an effective action, but you are likely to try to prevent or obstruct a bungling action by your teammate. By contrast, you are willing to help an opponent bungle, but your opponent is not likely to induce you to help him or her make an effective action (which, in effect, harms your chances of obtaining your goal).

The theory of cooperation and competition, then, goes on to make further predictions about different aspects of intrapersonal, interpersonal, intragroup, and intergroup processes from the predictions about substitutability, cathexis, and inducibility. Thus, assuming that the individual actions in a group are more frequently effective than bungling, among the predictions that follow from the theory are that *cooperative relations* (those in which the goals of the parties involved are predominately positive interdependence), as compared with competitive ones, show more of these positive characteristics:

- 1. *Effective communication is exhibited.* Ideas are verbalized, and group members are attentive to one another, accepting of the ideas of other members, and influenced by them. They have fewer difficulties in communicating with or understanding others.
- 2. *Friendliness, helpfulness, and lessened obstructiveness* are expressed in the discussions. Members also are more satisfied with the group and its solutions and favorably impressed by the contributions of the other group members. In addition, members of the cooperative groups rate themselves high in desire to win the respect of their colleagues.
- 3. The members of each group expect to be treated fairly and feel obliged to treat the others fairly.
- 4. Attempts to influence one another rely on persuasion and positive inducements.
- 5. *Coordination of effort, division of labor, orientation to task achievement, orderliness in discussion, and high productivity* are manifested in the cooperative groups (if the group task requires effective communication, coordination of effort, division of labor, or sharing of resources).

- 6. *Feeling of agreement with the ideas of others and a sense of basic similarity in beliefs and values, as well as confidence in one's own ideas and in the value that other members attach to those ideas*, are obtained in the cooperative groups.
- 7. *Recognizing and respecting the other by being responsive to the other's needs*.
- 8. *Willingness to enhance the other's power* (e.g., the knowledge, skills, resources) to accomplish the other's goals increases. As the other's capabilities are strengthened, you are strengthened; they are of value to you as well as to the other. Similarly, the other is enhanced from your enhancement and benefits from your growing capabilities and power.
- 9. *Defining conflicting interests as a mutual problem to be solved by collaborative effort* facilitates recognizing the legitimacy of each other's interests and the necessity to search for a solution responsive to the needs of all. It tends to limit rather than expand the scope of conflicting interests.

In contrast, a *competitive process* has the opposite effects:

- 1. Communication is impaired as the conflicting parties seek to gain advantage by misleading the other through use of false promises, ingratiation tactics, and disinformation. It is reduced and seen as futile as they recognize that they cannot trust one another's communications to be honest or informative.
- 2. Obstructiveness and lack of helpfulness lead to mutual negative attitudes and suspicion of one another's intentions. One's perceptions of the other tend to focus on the person's negative qualities and ignore the positive.
- 3. Fairness to the other is not valued. Each tries to exploit or harm the other to advantage themselves.
- 4. Attempts to influence the other often involve threat, coercion, or false promises.
- 5. The parties to the process are unable to divide their work, duplicating one another's efforts such that they become mirror images; if they do divide the work, they feel the need to check what the other is doing continuously.
- 6. The repeated experience of disagreement and critical rejection of ideas reduce confidence in oneself as well as the other.

- 7. The conflicting parties seek to enhance their own power and to reduce the power of the other. Any increase in the power of the other is seen as threatening to oneself.
- 8. The competitive process stimulates the view that the solution of a conflict can be imposed only by one side on the other, which in turn leads to using coercive tactics such as psychological as well as physical threats and violence. It tends to expand the scope of the issues in conflict as each side seeks superiority in power and legitimacy. The conflict becomes a power struggle or a matter of moral principle and is no longer confined to a specific issue at a given time and place. Escalating the conflict increases its motivational significance to the participants and may make a limited defeat less acceptable and more humiliating than a mutual disaster.

Constructive Competition

Competition can vary from destructive to constructive; unfair, unregulated competition at the destructive end; fair, regulated competition in between; and constructive competition at the positive end. In constructive competition, the losers as well as the winners gain. Thus, in a tennis match that takes the form of constructive competition, the winner suggests how the loser can improve, offers an opportunity for the loser to learn and practice skills, and makes the match an enjoyable or worthwhile experience for the loser. In constructive competition, winners see to it that losers are better off, or at least not worse off than they were before the competition.

The major difference, for example, between constructive controversy and competitive debate is that in the former, people discuss their differences with the objective of clarifying them and attempting to find a solution [p. 282 ↓] that integrates the best thoughts that emerge during the discussion, no matter who articulates them (see Johnson et al., 2006, for a fuller discussion). There is no winner and no loser; both win if during the controversy each party comes to deeper insights and enriched view of the matter that is initially in controversy. Constructive controversy is a process for constructively coping with the inevitable differences that people bring to cooperative interaction because it uses differences in understanding, perspective, knowledge, and worldview as valued resources. By contrast, in competitive contests or debates there

is usually a winner and a loser. The party judged to have “the best” – ideas, skills, knowledge, and so on – typically wins, while the other, who is judged to be less good, typically loses. Competition evaluates and ranks people based on their capacity for a particular task, rather than integrating various contributions.

I do not mean to suggest that competition produces no benefits. Competition is part of everyday life. Acquiring the skills necessary to compete effectively can be of considerable value. Moreover, competition in a cooperative, playful context can be fun. It enables one to enact and experience, in a nonserious setting, symbolic emotional dramas relating to victory and defeat, life and death, power and helplessness, dominance and submission; these dramas have deep personal and cultural roots. In addition, competition is a useful social mechanism for selecting those who are more able to perform the activities involved in the competition. Further, when no objective, criterion-referenced basis for measurement of performance exists, the relative performance of students affords a crude yardstick. Nevertheless, serious problems are associated with competition when it does not occur in a cooperative context and if it is not effectively regulated by fair rules (see Deutsch, 1973: 377–388, for a discussion of regulating competition).

Self-Destructive Tendencies Inherent in Cooperation

As I have indicated in my writings on cooperation and competition (Deutsch, 1973, 1985), there is a natural tendency for cooperation to break down as a result of the very social psychological processes – *substitutability*, *cathexis*, and *inducibility* – that are central to cooperation. Thus, *substitutability*, which enables the work of one cooperator to replace the work of another so that they don't have to duplicate one another's efforts, leads to specialization of function. Specialization of function, in turn, gives rise to specialized interest and to specialized terminology and language; the likely consequence is a deterioration of group unity as those with special interest compete for scarce resources and communicate in a language that is not fully shared. Similarly, *cathexis* of other group members (the development of personal favorable attitudes and bonds between members) can lead to in-group favoritism, clique formation,

nepotism, and so on. Here, the consequences are apt to be a weakening of overall group cohesion as cliques develop, a deterioration of cooperation with other groups as in-group favoritism grows, and a lessening of group effectiveness as a result of nepotism. *Inducibility*, the readiness to be influenced positively by other group members, can lead to excessive conformity with the view of others so that one no longer makes one's own independent, unique contribution to the group. The cooperative process, as a result, may be deprived of the creative contribution that can be made by each of its members, and also, those who suppress their individuality may feel inwardly alienated from themselves and their group despite their outer conformity. In addition, social loafing may occur in which some members shirk their responsibilities to the group and seek to obtain the benefits of group membership without offering the contributions they are able to make to it.

The Limitations of the Theory of Cooperation – Competition

My theory deals with pure simple situations of cooperation and competition, in which the interdependent parties each have only one goal and are equally interdependent. Of course, [p. 283 ↓] in real life this is rarely the case. In addition, the theory has not the precision and quantitative rigor and strong logical deductibility that ideally a theory should have. There are implicit “common sense” perceptual, cognitive, learning, and cultural assumptions within it that are necessary for its deductions. Like most theories in social psychology, it is not independent of other work on individual and social processes. In addition, my theory only considers the cooperation–competition dimension of social relations. As I have indicated elsewhere (Deutsch 1982, 1985, and in the subsequent section “Social relations and psychological orientations”), social relations differ not only in the cooperation–competition dimension but also in such other dimensions as: equality of power; task orientation versus social–emotional orientation; intimate versus formal; and importance of the relationship. The psychological orientation to a given social relation will be determined by the combined dimensions.

Despite the limitations of my theory, I consider it to be an important one because the dimension of cooperation–competition is one of the central variables of all social

relationships whether at the individual, group, or international level. (For a further discussion of the limitations and strengths of the theory see Johnson and Johnson, 2005: 326–342.)

A Theory of Conflict Resolution

After obtaining my PhD in the summer of 1948, I accepted a position at the Research Center for Human Relations (then at the New School) headed by Stuart Cook, which involved developing a comparative study of integrated and segregated interracial housing (Deutsch and Collins 1951). In 1949, the Center moved to New York University (NYU) where I initiated a program of research to develop insight into the conditions that affected the choice to cooperate or to compete. At NYU, I met Howard Raiffa, a scholar much interested in game theory and decision making (Luce and Raiffa, 1957), who introduced me to the prisoner's dilemma when I indicated my research interests. This led me to initiate research on the prisoner's dilemma and then on other mixed motive situations such as bargaining, negotiation, and conflict where there are typically a mixture of motivations to cooperate and to compete. As a result of doing research with such situations, we reformulated our questions from "What determines the choice to cooperate or compete?" to the conceptually similar but "sexier" "What determines whether a conflict will take a constructive or destructive course?" Our earlier research on the effects of cooperation and competition had indicated that a cooperative process was more likely to lead to constructive conflict resolution and a competitive process to a destructive resolution.

We did much research (Deutsch, 1973) in an attempt to find the answer. The results fell into a pattern I slowly began to grasp. They seemed explainable by an assumption I have immodestly labeled *Deutsch's Crude Law of Social Relations*:

The characteristic processes and effects elicited by a given type of social relationship also tend to elicit that type of social relationship; and a typical effect tends to induce the other typical effects of the relationship.

Thus, cooperation induces and is induced by perceived similarity in beliefs and attitudes; readiness to be helpful; openness in communication; trusting and friendly attitudes; fair treatment; sensitivity to common interests and de-emphasis of opposed interests; orientation toward enhancing mutual power rather than power differences, and so on. Similarly, competition induces and is induced by use of the tactics of coercion, threat, or deception; attempts to enhance the power differences between oneself and the other; poor communication; minimization of the awareness of similarities in values and increased sensitivity to opposed interests; suspicious and hostile attitudes; unfair treatment; and so on.

In other words, if one has systematic knowledge of the effects of cooperative and competitive processes, one has systematic knowledge of the conditions that typically [p. 284 ↓] give rise to such processes and, by extension, to the conditions that affect whether a conflict takes a constructive or destructive course. My early theory of cooperation and competition is a theory of the effects of cooperative and competitive processes. Hence, from the Crude Law of Social Relations, it follows that this theory brings insight into the conditions that give rise to cooperative and competitive processes.

This law is certainly crude. It expresses surface similarities between effects and causes; the basic relationships are genotypical rather than phenotypical. The surface effects of cooperation and competition are due to the underlying type of interdependence (positive or negative) and type of action (effective or bungling), the basic social psychological processes involved in the theory (substitutability, cathexis, and inducibility), and the cultural or social medium and situational context in which these processes are expressed. Thus, how a positive attitude is expressed in an effective, positively interdependent relationship depends on what is appropriate to the cultural or social medium and situational context; presumably one would not seek to express it in a way that is humiliating or embarrassing or likely to be experienced negatively by one's partner.

Similarly, the effectiveness of any typical effect of cooperation or competition as an initiating or inducing condition of a cooperative or competitive process is not due to its phenotype but rather to the inferred genotype of the type of interdependence and types of action. Thus, in most social media and social contexts, perceived similarity in basic

values is highly suggestive of the possibility of a positive linkage between oneself and the other. However, we are likely to see ourselves as negatively linked in a context that leads each of us to recognize that similarities in values impel seeking something that is in scarce supply and available for only one of us. Also, it is evident that although threats are mostly perceived in a way that suggests a negative linkage, any threat perceived as intended to compel you to do something that is good for you or that you feel you should do is apt to be suggestive of a positive linkage.

Although the law is crude, my impression is that it is reasonably accurate; phenotypes often indicate the underlying genotypes. Moreover, it is a synthesizing principle, which integrates and summarizes a wide range of social psychological phenomena. One can integrate much of the literature on the determinants of positive and negative attitudes in terms of the other associated effects of cooperation and competition. Thus, positive attitudes result from perceptions of similarity, open communication, and so on. Similarly, many of the determinants of effective communication can be linked to the other typical effects of cooperation or competition, such as positive attitude and power sharing.

In brief, the *theory of conflict resolution* equates a constructive process of conflict resolution with an effective cooperative problem-solving process in which the conflict is the mutual problem to be resolved cooperatively. It also equates a destructive process of conflict resolution with a competitive process in which the conflicting parties are involved in a competition or struggle to determine who wins and who loses; often, the outcome of the struggle is a loss for both parties. The theory further indicates that a cooperative-constructive process of conflict resolution is fostered by the typical effects of cooperation and a competitive-destructive process by the typical effects of competition. The theory of cooperation and competition outlined in the beginning of this part is a well-verified theory of the effects of cooperation and competition and thus allows insight into what can give rise to a constructive or destructive process.

Limitations and Strengths of the Theory of Conflict Resolution

As indicated above, the theory is based upon two key elements: the theory of the effects of cooperation and competition processes and the Crude Law of Social Relations. The limitation and strengths of these two elements [p. 285 ↓] lead to the limitations and strengths of the theory of conflict resolution. The main difference between the strengths of the two elements is that the theory of cooperation–competition has generated much supportive research; apart from the research reported in Deutsch (1973, 1985), little research on the Crude Law has yet been done. Both theoretical elements deal with processes central to social psychology and social life.

Recently, two papers have been prepared which present formal models concerned with conflict: “Dynamics of two-actor cooperation – conflict models” (Liebovitch et al., 2008) and “From crude law to precise formalism: identifying the essence of conflict intractability” (Nowak et al., 2008). They are respectively relevant to the two elements of the theory of conflict resolution.

Distributive Justice

My theorizing and research in this area has mainly focused on two central questions: What are the effects of different principles of distributive justice? And what leads to preference for one principle or another? My book, *Distributive Justice* (Deutsch, 1985) presents much of the relevant work by my students and myself.

My work on the social psychology of justice was initiated by an invitation from Melvin Lerner, a social psychologist who has made many important contributions to this area. Early in 1972, he invited me to write a paper for a Conference on Injustice in North America (Deutsch, 1974, 1985). In preparation for the 1972 conference, I read widely – delving into the literature of the moral and legal philosophers, sociologists, and political scientists, and the relevant work of social psychologists. The more I read, the more dissatisfied I became with the existing literature in social psychology; it seemed too

narrowly focused, too parochial, and too unwittingly reflective of the dominant Western ideology. The focus was limited mainly to how subjects in laboratory experiments attempted to restore their psychological equilibrium after experiencing or observing an inequity. There was little research on such topics as the conditions necessary for awakening the sense of injustice, procedural justice, retributive justice, and so on. The emphasis of equity theorists on “proportionality” as the sole canon of distributive justice suggested that they were neglecting other distributive principles, such as “equal share to all” or “to each according to his need,” which have been rallying slogans for different political ideologies (Deutsch, 1975). Beyond this, the economic and market orientation of equity theory appeared to reflect, unwittingly, the implicit assumption in much of current Western ideology that economic “rationality” and economic values should pervade all social life and are appropriate in noneconomic social relations (e.g., between lovers, between parent and child).

I note that as I became more involved in this area and started to think about my research – past and present – in the context of “justice,” I found myself in the position of the bourgeois gentleman of Moliere’s play who was delighted to learn that he had been speaking prose all the time. I was delighted to recognize that my research under other labels could be labeled as “justice” research quite properly. Thus, my early study of the effects of cooperation and competition upon group processes (Deutsch, 1949b) could be considered a study of the consequences of two contrasting distributive values (“rewarding group members equally” or “rewarding them in terms of their relative rank in their contributions to group performance”). Similarly, our many studies of conflict and bargaining (Deutsch, 1973) are centrally related to the social psychology of justice. They were focused on the important questions: Under what conditions are people with conflicting interests able to work out an agreement (i.e., a system of justice defining what each shall give and receive in the transaction between them) that is stable and mutually satisfying?

In *Distributive Justice*, I concentrated on four distributive principles: *winner-takes-all*, *equity* (proportionality), *equality*, and *need*. I was interested in what effects these principles would have in tasks where interdependent work was neither required nor possible, and in tasks where interdependent work was necessary.

The principles were described as follows:

- *Winner-takes-all*: Under this system, whoever performs the task best in the group wins all the money the group is paid.
- *Proportionality*: Under this system, each person is rewarded in proportion to his or her contribution to the group score. In other words, the person who contributes 50 percent of the group's total output will get 50 percent of the money to be distributed within the group; a person who contributes 10 percent would get 10 percent, and so on.
- *Equality*: Under this system, each person in the group will get an equal share of the money to be distributed within the group.
- *Need*: Under the need distribution system, each group member will be rewarded according to the need expressed on a biographical data sheet. In other words, the person who needs the money most will get proportionately more money; the person who needs the money least will get the least amount of money.

The theory of cooperation and competition and the Crude Law were employed to develop hypotheses about both the effects and the choice of the different principles. With the assumption that our college student subjects were not alienated (from themselves, work, or the experiment) and would work as well as they could, we predicted that there would be no significant differences in the productivity of the subjects when no interdependent work was required or possible. However, we also predicted that the different principles would elicit different attitudes toward the other group members; the more cooperative principles (*equality* and *need*) would elicit more favorable attitudes than the more competitive principles (*winner-takes-all* and *equity*).

In the tasks where interdependent work was necessary, we predicted that the results would be similar to those obtained in the earlier research on cooperation – competition; higher productivity and more favorable attitudes when *cooperative* rather than *competitive* distributive principles were employed. The results of our various experimental studies (see Deutsch, 1985) were supportive of our hypotheses.

With regard to the choice among the different principles, we employed The Crude Law to make predictions. In my paper, "Interdependence and psychological orientation" (Deutsch 1982, 1985), I developed the idea that different social relations require different psychological orientations (see the section "Social relations and

psychological orientations,” for further elaboration). Based upon prior research on the effects of cooperation and competition as well as research on the effects of the different distributive justice principles, we characterized two psychological orientations: *solidarity* and *economic*. The *solidarity* orientation is congruent with a relationship that is cooperative, equal, social–emotional, and informal, while the *economic* fits a relationship that is competitive, equal, task-oriented, and formal. A *solidarity* orientation is defined by a sense of positive bonding and positive feelings toward and from the others; more reliance on empathy and intuition in understanding the others; and an awareness of a mutual obligation to be helpful to one another. An *economic* orientation is characterized by detachment, an objective–analytical perspective, a utilitarian self-interest, and expectation that the others have a similar interest.

In Chapter 11 of *Distributive Justice* (Deutch, 1985), some research bearing upon these ideas is presented. The research is supportive, but only a few studies were conducted. I do not consider that the evidence is strong, but I believe that the underlying ideas are. In the following text, I elaborate on some of these ideas.

Social Relations and Psychological Orientation

A number of years ago, I was doing a study of marital couples and I wanted to develop a way of characterizing the nature of the couple relationship. With the help of Myron Wish [p. 287 ↓] (Wish et al., 1976), we developed a method of doing so. In the course of doing so, we identified what we considered to be several of the basic dimensions of social relations: cooperation–competition; power distribution; task-orientated versus social–emotional; formal versus informal; degree of importance. Some of these are similar to those described by other investigators.

In terms of these dimensions, “friends” would be generally considered to be cooperative, of equal power, in a social–emotional, and informal relationship of considerable importance. In contrast, the relationship between a police officer and a thief might be viewed as competitive, unequal power, task-oriented, formal, and of moderate importance.

My next thought was that to act appropriately in a given type of social relation one must have an appropriate psychological orientation to that relationship: one's psychological orientation must "fit" the social relation. For example, my psychological orientation when I am negotiating the price of a car with a used-car salesman will be rather different than when I am playing with my six-year-old grandson. Different types of social relations will induce different types of psychological orientations, and, according to my "Crude Law," different types of psychological orientations will induce different types of social relations.

The Nature of Psychological Orientations

In my current view, a psychological orientation consists of four highly interdependent elements: a cognitive orientation, a motivational orientation, a moral orientation, and an action orientation. In my prior publications (Deutsch 1982, 1985), action orientation was not included.

Cognitive Orientations

In recent years, scholars in a number of different disciplines – cognitive psychology, social psychology, sociology, linguistics, anthropology, and artificial intelligence – have utilized such terms as *schema*, *script*, and *frame* to refer to the *structures of expectations* that help orient the individual cognitively to the situation confronting her. I employ the term *cognitive orientation* as being essentially the same. In the view being presented here, the person's cognitive orientation to his situation is only one aspect of his psychological orientation to a social relationship.

Underlying the concepts of schema, script, and frame is the shared view that people approach their social world actively, with structured expectations about themselves and their social environments that reflect their organized beliefs about different social situations and different people. Our structured expectations make it possible for us to interpret and respond quickly to what is going on in specific situations. If our expectations lead us to inappropriate interpretations and responses, then they are likely to be revised on the basis of our experiences in the situation. Or if the circumstance

confronting us is sufficiently malleable, our interpretations and responses to it may help to shape its form.

It is important for the participants in a particular social relationship to know “what's going on here” – to know the actors; the roles they are to perform; and the relations among the different roles, the props and settings, the scenes, and the themes of the social interaction. However, everyday social relations are rarely as completely specified by well-articulated scripts as the social interaction in a play in the traditional theater; ordinary social interactions have more the qualities of improvisational theater in which only the nature of the characters involved in the situation is well-specified and the characters are largely free to develop the detail of the skeletonized script as they interact with one another.

The improvisational nature of most social relations – the fact that given types of social relations occur in widely different contexts and with many different kinds of actors – makes it likely that relatively abstract or generalized cognitive orientations will develop from the different types of social relations. I assume that people are implicit social psychological theorists and, as a result of their experience, have developed cognitive [p. 288 ↓] schemas of the different types of social relations, though usually not articulated, that are similar to those articulated by theorists in social psychology and the other social sciences. Undoubtedly, at this early stage of the development of social science theory, the unarticulated conceptions of the average person are apt to be more sophisticated than the articulated ones of the social scientists.

Motivational Orientations

Just as different cognitive orientations are associated with the different types of social relations, so also are different motivational orientations. A motivational orientation toward a given social relationship orients one to the possibilities of gratification or frustration of certain types of needs in the given relationship. To the cognitive characterization of the relationship, the motivational orientation adds the personal, subjective features arising from one's situationally relevant motives or need-dispositions.

The motivational orientation gives rise to the cathexis of certain regions of the cognitive landscape, making them positively or negatively valent, and highlights the pathways to and from valent regions. It gives the cognitive map a dynamic character. It predisposes one to certain kinds of fantasies (or nightmares) and to certain kinds of emotions. It orients one to such questions as, "What is to be valued in this relationship?" and "What do I want here and how do I get it?"

Moral Orientations²

A moral orientation toward a given social relationship direct one to the mutual obligations, rights, and entitlements of the people involved in the given relationship. It adds an "ought," "should," or obligatory quality to a psychological orientation. The moral orientation implies that one experiences one's relationship not only from a personal perspective but also from a social perspective that includes the perspective of the others in the relationship. A moral orientation makes the experience of injustice more than a frustrating, personal experience (Deutsch et al., 1978). Not only is one personally affected, so are the other participants in the relationship, because its value underpinnings are being undermined. The various participants in a relationship have the mutual obligation to respect and protect the framework of social norms that define what is to be considered as fair or unfair in the interactions and outcomes of the participants. One can expect that the moral orientation, and hence what is considered fair, will differ in different types of social relations.

Action Orientation

Action orientations refer to the kinds of behavior which are viewed as appropriate in a given type of social relationship. Different cultures often have different views as to what is appropriate behavior in a given social relationship. Thus, if I felt very pleased with the outcome of my negotiations with the used-car salesman (I got a very good price), it would be inappropriate behavior to express my pleasure by kissing him.

The moral component that exists in all social relations has largely been neglected in social psychological theorizing. I suggest that the *moral orientation*, what is perceived to be just or unjust, will vary in different types of social relations. Let me illustrate the moral component of several different types of social relationships.

Equality–Inequality

There are a number of different moral orientations connected with equality and inequality: other features of the relationship, in addition to the distribution of power within it, will determine the nature of the moral orientation that will be elicited. Thus, in a cooperative, equal relationship one would expect an egalitarian relationship. In a cooperative, unequal relationship, the moral orientation obligates the more powerful person to employ his power in such a way as to benefit the less powerful one, not merely himself. In such a relationship, the less powerful one has the obligation to show appreciation, to defer to, and honor the more powerful person. These obligations may be [p. 289 ↓] rather specific and limited if the relationship is task-oriented or they may be diffuse and general if the relationship is a social-emotional one.

In an equal, competitive relationship, one's moral orientation is toward the value of initial equality among the competitors and the subsequent striving to achieve superiority over the others. This orientation favors equal opportunity but not equal outcomes: the competitors start the contest with equal chances to win, but some win and some lose. In an unequal, competitive relationship the moral orientations of the strong and the weak support an exploitative relationship. The strong are likely to adopt the view that the rich and powerful are biologically and hence morally superior; they have achieved their superior positions as a result of natural selection; it would be against nature to interfere with the inequality and suffering of the poor and weak; and it is the manifest destiny of superior people to lead inferior people. In an unequal, competitive relationship, the weak are apt to *identify with the aggressor* (Freud, 1937) and adopt the moral orientation of the more powerful and to feel that their inferior outcomes are deserved. Or, they may feel victimized. If so, they may either develop a revolutionary moral orientation directed toward changing the nature of the existing relationship or they may develop the moral orientation of being a victim. The latter orientation seeks to obtain secondary

gratification from being morally superior to the victimizer. “It's better to be sinned against than to sin;” “The meek shall inherit the earth.”

Task versus Social–Emotional Relations

The moral orientation in a task-oriented relationship is that of utilitarianism. Its root value is maximization: people should try to get the most out of a situation. Good is viewed as essentially quantitative, as something that can be increased or decreased without limit (Diesing, 1962: 35). A second element in this moral orientation is the means-end schema, in which efficient allocation of means to achieve alternative ends becomes a salient value. A third element is impartiality in the comparison of means, so that means can be compared on the basis of their merit in achieving given ends rather than on the basis of considerations irrelevant to the means-end relationship. In Parsonian terms, the moral orientation in task-oriented relations is characterized by the values of universalism, affective neutrality, and achievement. In contrast, the moral orientation of social–emotional relations is characterized by the values of particularism, affectivity, and ascription (Diesing, 1962: 90). Obligations to other people in a social–emotional relationship are based on their particular relationship to oneself rather than on general principles: they are strongest when relations are close and weakest when relations are distant. In a task-oriented relation, one strives to detach oneself from the objects of one's actions, to treat them all as equal, separate, interchangeable entities. In a social–emotional relationship, one is the focal point of myriad relationships that one strives to maintain and extend, since action takes place only within relationships (Diesing, 1969: 91). Ascription is the opposite of the achievement value: it means that one's action and obligations toward people spring solely from their relationship to oneself rather than as a response to something they have done.

Some Potential Research

From the Crude Law, it follows that the causal arrow connecting psychological orientations and types of social relations is bidirectional: a psychological orientation can induce or be induced by a given type of social relation. Here, I would go further

and indicate that the cognitive, motivational, and moral components of a psychological orientation can each induce one another – hence, they are likely to be found together – and each of the components can induce or be induced by a given type of social relation. The foregoing assumptions proliferate into a great number of testable, specific hypotheses that would predict a two-way causal arrow between specific modes of thought, [p. 290 ↓] and specific types of social relations. Thus, a bureaucratic social situation will tend to induce obsessive-compulsive modes of thought and obsessive-compulsive modes of thought will tend to “bureaucratize” a social relationship. They would also predict that a competitive social relationship will tend to increase the psychological weight or importance of the difference in values between oneself and one's competitors, whereas a cooperative relationship will tend to increase the psychological importance of the similarities in values between oneself and one's fellow cooperators. We would also hypothesize that a tendency to accentuate the difference in values between oneself and others is apt to induce a competitive relationship, whereas a tendency to accentuate the similarities is likely to induce a cooperative relationship. Further, it can be predicted that different principles of distributive justice will be associated with different types of social relations: a fraternal relationship will be connected with the principle of equality, a caring relationship with the principle of need, a hierarchical organization with the principle of equity, and a power struggle with the principle of winner takes all.

Limitations of Social Relations and Psychological Orientations

This work, more fully presented in Deutsch (1982, 1985), is a sketch of some important theoretical ideas. It needs much more theoretical development and much more research.

Social Implications of My Work

I have always considered my contributions to psychology as being theoretical and myself as someone who developed theoretical ideas and did research related to theory.

However, my mentor, Kurt Lewin taught his students that 'there is nothing as practical as a good theory'. While I did not anticipate the practical applications of my work, I believe that it has had some important ones. I will describe a few.

First, my dissertation study, a theoretical and experimental study of the effects of cooperation and competition upon group process (Deutsch, 1949a, 1949b) was done in the context of small experimental classes in an undergraduate psychology course I was teaching at MIT. Although the guiding image underlying my study related to issues of war and peace, to whether the then-recently created UN Security Council would function cooperatively or competitively, my experiment involved the creation of cooperative and competitive small classrooms of five students. I published a paper in an education journal on the educational implications of my theory and research. However, it was David W. Johnson – a former doctoral student of mine – who systematically developed these ideas into a pedagogy of cooperative learning and helped many teachers and school systems through the world to adopt this approach to education. It has also been widely applied in industry (see Johnson and Johnson, 2005).

Second, I directed a study on interracial housing (Deutsch and Collins, 1951) which compared the behavioral and attitudinal effects of living in public housing where the white and black residents were integrated (living in the same building) or segregated (living in separate buildings) within the housing project. The integrated housing was in New York City; the segregated housing in Newark. The results of this study played a role in changing the Newark Public Housing from a policy of segregating to integrating the races in their housing projects. I quote from a statement made by the Director of the Newark Housing Projects (from back cover of *Interracial Housing*):

A new policy ... provides that henceforth all apartments are to be allocated on a basis of need, regardless of race, religion, and color ... In large measure, this change in fundamental policy reflects the impact of the study reported in this book.

Deutsch and Collins, 1951: back cover

The study not only affected policy in Newark, it played a role in changing policies of the US Public Housing Authority which provided [p. 291 ↓] some of the financing

for local housing authorities. Additionally, it was a small part of the material that a SPSSI Committee (which included Kenneth Clark, Isadore Chein, and me) prepared for the lawyers who successfully petitioned the US Supreme Court in *Brown v. Board of Education* to end racial segregation in publicly supported schools.

Third, my work on conflict resolution, with the help of many former students and many other scholars, helped to stimulate the development of the field of conflict resolution studies. The basic query underlying our theoretical and research work on conflict ("What determines whether a conflict will take a constructive or destructive course?") has direct relevance to conflicts in the real world. I and many of my former students have applied our theoretical work on conflict to such diverse conflicts as marital conflict, intergroup and ethnic conflict, industrial conflict, educational conflict, international conflict, reconciliation after destructive conflict, and so on. The applications have taken various forms; analytical writing, education, workshops with practitioners, mediation, and consultation to the conflicting parties. Such students as Jeffrey Rubin, Roy Lewicki, David Johnson, Michelle Fine, Harvey Hornstein, Madelaine Heilman, Barbara Bunker, Kenneth Kressel, Susan Opotow, Janice Steil, Peter Coleman, Eric Marcus, Ken Sole, Adrienne Asch, and many others have made important, original contributions to the development of practice as well as theory in this area. In addition, the Center that I founded at Teachers College, the International Center for Cooperation and Conflict Resolution (ICCCR), has helped to stimulate the development of conflict resolution and mediation programs in many schools.

For me, one of the applications of my work is an unusually important one. It occurred in Poland where two outstanding psychologists, Janusz Reykowski and Janusz Grzelak, applied some of my ideas during the negotiations between the Communist government and Solidarity which lead to a *peaceful* transfer of governmental power from the Communist Party to Solidarity in 1989. Reykowski was a leading figure in the Communist Party and Grzelak was a very important influence in the Solidarity movement. Each has indicated that my work influenced him considerably in their approach to the negotiations which facilitated a constructive resolution of the negotiations. In footnote #3, I quote from some remarks made by Professor Reykowski at a Conference in Intractable Conflict held in Poland in the fall of 2006.³ In footnote #4,

I quote from statements made by Professor Grzelak about his role in the negotiations and about my influence in an email sent to Lan Bui-Wrzosinska, and forwarded to me.⁴

The social effects of my work in the areas of cooperation–competition, inter-racial housing, and conflict resolution have been notable. I do not yet have a clear picture of the direct social impact of my work in the area of social justice. Three papers of mine appear to have had a considerable impact on the social psychological study of justice and to have been widely used in classrooms: “Equity, equality, and need” (Deutsch, 1973), “Awakening the sense of injustice” (Deutsch, 1974, 1985), and “A framework for thinking about oppression and its change” (Deutsch, 2006).

Let me conclude this section by stating: I did not foresee many of the applications of my theoretical and empirical work. Like throwing a pebble into water, the ripples of one's theoretical work are hard to predict in advance.

Conclusion

When I taught a course on theories in social psychology, I suggested to students that there were two types of theorists: *grandiose* and *picayune*. The *grandiose* theorists generalize their ideas widely and freely, the *picayune* keep their generalizations very close to their data. I consider myself to fall into the *grandiose* category. I have generalized my ideas so that they are relevant not only to the [p. 292 ↓] individual as the social interactor but also to the other types such as groups and nations. I have done this because I think my ideas deal with basic social processes. Clearly, my ideas are not fully baked and they need many more and different ingredients to deal with the different types of social actors. It is my hope that others will finish the baking of these ideas.

Notes

1 There was much open antisemitism in the United States during the 1920s, 1930s, and 1940s, and I experienced some of it directly during this period.

2 In my view, social psychologists have unduly neglected the moral aspect of every social relation. In my study of conflict, I was alerted to how issues of justice–injustice often played a central role in conflict. This led me to reflect on the difference between the experiences of injustice and frustration (Deutsch et al., 1978) which, in turn, led me to think about the moral norms in different social relations. Also, during the 1960s at Columbia, some students flouted the moral or social norms in an attempt to bring about social change; for example, by appearing naked in a classroom. Additionally, the work of Goffman (1959) was suggestive. By the “moral aspect of every situation,” I refer to the social norms which define appropriate and inappropriate behavior and the mutual obligations which enable the social relation to exist.

3 Professor Reykowski's remarks at the Conference on Intractable Conflict in Poland, 2006:

Most of the great ideas produced by psychologists are appreciated for their intellectual value rather than for their consequences for practical life. There are, however, some exceptions. For me one such exception is Morton Deutsch because there are good reasons to claim that his theories went beyond academia and have had an impact on some large scale social processes – that took place a thousand miles from Morton's home place. In fact, they took place in Poland in the middle of the eighties. It was a period of time when Poland ... was in the state of deep crisis. “Solidarity” – the massive democratic movement had been crushed during Martial Law (introduced in Poland in December 1981) and the country was overwhelmed by a major political and social conflict ... As a psychologist, I was especially interested in analyzing the psychological factors that contributed to the development of the conflict situation and in possible psychological remedies. That was why I focused on Morton Deutsch's *The Resolution of Conflict* (Yale University, 1973) that I received from him some time ago. And now it seemed to offer the insight that I needed. The major theses of the book ... provided excellent conceptual instruments for description of the Polish situation and were a very good source of ideas for developing proposals how to deal with it. I wrote an article in the major, very influential, Polish weekly magazine (*Polityka*) – widely read by

intelligentsia and members of the establishment – presenting Morton's theory and indicating how it could be applied to the Polish context. The approach met with an attack from both sides ...

Unlike earlier time, the attack in the official party newspaper was not a political death sentence for its author. To the contrary, I was allowed to respond to the criticism in the same newspaper and attacking my opponents I could further describe the concepts of destructive [and constructive] conflict and their importance for understanding the Polish situation. I have some reasons to believe that this exchange and my further activities along this line had some impact on members of the ruling elite in Poland. A few years later, when the ruling party came to the conclusion that the policy of accommodation with Solidarity is a necessary step for solving the Polish conflict, I was called upon to help in execution of this policy. The most important first step of this new policy was the Round Table negotiations between Government and Solidarity. As a result of these negotiations the partially free election took place in Poland and following that the new government held by Tadeusz Mazowiecki a leading Solidarity figure was introduced. In other words, the starting point for a series of events that led to dissolution of the so called Soviet Bloc...

The most important ... [Round Table] was the political table because there was a place where the main political changes were formulated and negotiated. I was a co-chair of the political table ... It is not a place for detailed description of the negotiation. I would like to conclude that Morton Deutsch's theory of destructive conflict had not only an important place in psychological science but also has some place in the history of social change in Europe.

4 Quotation from an e-mail by Professor Grzelak:

I was one of the two vice-chairmen of one of the Round Table workgroups (the chairman was Prof. Henryk Samsonowicz) – concerning education and science. I was also asked, as an expert, to

participate in the informal, although not secret, talks in smaller groups. Several of them prepared the most important decisions. I took part in probably all of them, mostly as the main negotiator ... Due to the [p. 293 ↓] personal interests in psychology and thanks to 2.5 years spent in the United States, I did know a number of works concerning conflict of interest. Among them most important were those by Kelley, Rapaport and Deutsch. I mentioned the Author whom I owe especially, last. His theoretical (On cooperation and competition, 1949; On the resolution of conflict, 1973; and many more) and empirical works ... helped understand the role of trust, the role of orientation in an interaction, the power of power and the weakness of power, when it is used in conflict management. Theory is the most important but equally important is the "spirit" of the theory. What Morton said and says is filled with respect for people, for their subjectivity, it's a constant search for resolutions both just and satisfying ... My fascination in conflict resolution began with reading Deutsch's works long before the downfall of communism and stayed alive long after the downfall of communism.

References

- Barnard, C.I. (1938) *The Functions of the Executive*. Cambridge, MA: Harvard University Press.
- Deutsch, M. A theory of cooperation and competition . *Human Relations* , 2, 129–151. (1949a) <http://dx.doi.org/10.1177/001872674900200204>
- Deutsch, M. An experimental study of the effect of cooperation and competition upon group processes . *Human Relations* , 2, 199–231. (1949b) <http://dx.doi.org/10.1177/001872674900200301>
- Deutsch, M. Field theory in social psychology . In G. Lindzey, ed. and E. Aronson (eds), *The Handbook of Social Psychology* , 1, 412–487. (1954)

Deutsch, M. (1973) *The Resolution of Conflict: Constructive and Destructive Processes*. New Haven, CT: Yale University Press.

Deutsch, M. (1974) Awakening the sense of injustice . In M. Ross, ed. and M. Lerner (eds), *The Quest for Justice* , pp. 19–42. Canada: Holt, Reinhart, and Winston.

Deutsch, M. Equity, equality, and need . *Journal of Social Issues* , 31, 137–149. (1975) <http://dx.doi.org/10.1111/j.1540-4560.1975.tb01000.x>

Deutsch, M. (1982) Interdependence and psychological orientation . In V. Derlega, ed. and J.L. Grzelak (eds), *Cooperation and Helping Behavior* , pp. 16–41. New York: Academic Press.

Deutsch, M. (1985) *Distributive Justice, A Social-psychological Perspective* . New Haven, CT: Yale University Press.

Deutsch, M. Kurt Lewin: The tough-minded and tender-hearted scientist . *Journal of Social Issues* , 48, 31–43. (1992) <http://dx.doi.org/10.1111/j.1540-4560.1992.tb00881.x>

Deutsch, M. A framework for thinking about oppression and its change . *Social Justice Research* , 19, 7–41. (2006) <http://dx.doi.org/10.1007/s11211-006-9998-3>

Deutsch, M., ed. , Coleman, P.T., ed. and Marcus, E.T. (eds) (2006) *The Handbook of Conflict Resolution: Theory and Practice* , 2nd Edition. San Francisco: Jossey-Bass.

Deutsch, M. and Collins, M.E. (1951) *Interracial Housing: A Psychological Evaluation of a Social Experiment* . Minneapolis: University of Minnesota Press.

Deutsch, M., Steil, J. and Tuchman, B. An exploratory study of the meanings of injustice and frustration . *Personality and Social Psychology Bulletin* , 4, 393–398. (1978) <http://dx.doi.org/10.1177/014616727800400307>

Diesing, P. (1962) *Reason in Society* . Urbana: University of Illinois Press.

Freud, S. (1937) *The Ego and the Mechanisms of Defense* . London: Hogarth.

Goffman, E. (1959) *The Presentation of Self in Everyday Life* . Garden City, NY: Doubleday Anchor Books.

Johnson, D.W. and Johnson, R.T. New developments in social interdependence theory . *Psychology Monographs* , 131, 285–360. (2005)

Johnson, D.W., Johnson, R.T. and Tjosvold (2006) Constructive controversy: The value of intellectual opposition . In M. Deutsch, ed. , P. Coleman, ed. and E.C. Marcus (eds), *The Handbook of Conflict Resolution: Theory and Practice* , 2nd Edition. San Francisco: Jossey-Bass.

Koffka, K. (1935) *Principles of Gestalt Psychology* . New York: Harcourt, Brace.

Lewin, K. (1935) *Dynamic Theory of Personality* . New York: McGraw-Hill.

Lewin, K. (1936) *Principles of Topological Psychology* . New York: McGraw-Hill. <http://dx.doi.org/10.1037/10019-000>

Lewin, K. The conceptual representation and measurement of psychological forces . *Contributions to Psychological Theory* , 1 (4) (1938)

Lewin, K. (1948) *Resolving Social Conflicts* . New York: Harper and Brothers.

Liebovitch, L.S., Vallacher, R. and Michael, J. Dynamics of two-actor cooperation-competition models . *Peace and Conflict: Journal of Peace Psychology* , 16, 175–188. (2008). <http://dx.doi.org/10.1080/10781911003691625>

Luce, R.D. and Raiffa, H. (1937) *Games and Decisions* . New York: Wiley.

Maller, J.B. (1929) Cooperation and competition: An experimental study in motivation . *Teachers College Contribution to Education*, 384 .

May, M.A. and Doob, L.W. (1937) Cooperation and competition . *Social Science Research Council Bulletin*, 125 .

Mead, G.H. (1934) *Mind, Self and Society* . Chicago: University of Chicago Press.

Mead, M. (1937) Cooperation and Competition Among Primitive People . New York: McGraw-Hill. <http://dx.doi.org/10.1037/13891-000>

Nowak, A., Deutsch, M., Bartowski, W. and Solomon, S. From Crude Law to precise formulation: Modeling the path to destructiveness in conflict . Please and Conflict: Journal of Peace Psychology , 16, 189–210. (2008) <http://dx.doi.org/10.1080/10781911003691617>

Wish, M., Deutsch, M. and Kaplan, S. Perceived dimensions of interpersonal relations . Journal of Personality and Social Psychology , 33, 400–420. (1976) <http://dx.doi.org/10.1037/0022-3514.33.4.409>

<http://dx.doi.org/10.4135/9781446249222.n40>

Handbook of Theories of Social Psychology

The Focus Theory of Normative Conduct

Contributors: Robert B. Cialdini

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "The Focus Theory of Normative Conduct"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n41>

Print pages: 295-313

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n41>

[p. 295 ↓]

Chapter 41: The Focus Theory of Normative Conduct

Abstract

The focus theory of normative conduct offers a way to make sense of the mixed support in the behavioral sciences for the role of social norms in human behavior. On the basis of the results of subsequent theory-relevant research, it appears that norms do have a considerable impact on behavior, but that the force and form of that impact can only be usefully understood through certain conceptual elements that have not been traditionally or rigorously applied. That is, to predict properly the likelihood of norm-consistent action requires that one must (a) separate two types of norms that at times act antagonistically in a situation – injunctive norms (what most others approve/disapprove) and descriptive norms (what most others do) – and (b) focus individuals' attention principally on the type of norm desired to operate. General support for the theory has emerged via research conducted in a variety of naturally occurring settings (e.g., a parking garage, amusement park, suburban neighborhood, upscale hotel), while employing a variety of communication vehicles (e.g., handbills, park signage, door-hangers, public service announcements), which generated significant change in a variety of environment-relevant activities (e.g., littering, recycling, energy usage, environmental crime).

Introduction

I was raised in an entirely Italian family, in a predominantly Polish neighborhood, in a historically German city (Milwaukee), in an otherwise rural state. I often ascribe my interest in the social influence process to an early recognition that the groups populating those settings had to be approached somewhat differently in order to obtain their assent, sometimes to the identical request. It also struck me early on that one

reason for this complication was that the social norms – the characteristic tendencies and codes of conduct of the groups – differed. Therefore, if I wanted to maximize compliance with a request from a member of one or another of these groups, it would be wise to take into account the dominant norms of that particular unit.

This vague understanding that social norms both vary and make a difference had little impact on my choice of topics to study during graduate training and in the initial stages of my research career; indeed, the chosen topics in those days rarely involved normative issues, [p. 296 ↓] and I rarely thought about the social influence process I was studying at the time in normative terms. That began to change, though, when a few of my investigations encountered and even incorporated concepts that could be understood in normative terms: reciprocity, consistency, and social responsibility. I became aware that each of these normative concepts varied (within and between cultures) and made a notable difference in influencing human behavior, as some of my subsequent research has tried to document (Bator and Cialdini, 2006; Cialdini et al., 1995, 1999; Petrova et al., 2007).

However, the major impetus for a *theory* of normative conduct came not from a comfortable state of affairs – the easy fit I could see between social norms and social influence – but from an intriguingly discomforting one: certain thinkers whose judgment in theoretical matters I trusted greatly felt very differently about the usefulness of social norms as an explanatory and predictive concept than did other thinkers whose judgment I trusted just as much. On the one hand were those who saw the concept as central to a proper understanding of human social behavior (e.g., Berkowitz, 1972; Fishbein and Ajzen, 1975; Pepitone, 1976; Triandis, 1977). On the other hand were those who saw little value in the concept, viewing it as vague, overly general, often contradictory, and ill-suited to empirical tests (e.g., Darley and Latané, 1970; Krebs, 1970; Krebs and Miller, 1985; Marini, 1984).

By this time, I was well into my career and had come to recognize that a battle between heavyweights in an important theoretical arena is more than merely interesting or beteworthy. It presents an uncommon opportunity for anyone willing to try to settle the argument, an opportunity to do more than just take a side but perhaps to make a genuine contribution by offering a novel way of thinking about the topic that resolves the conflict. With the focus theory of normative conduct and with the invaluable

collaboration of excellent coworkers, it has been my hope to make the best of one such recognized opportunity.

Early Years: Profiting from the Complaints

Although I hadn't paid much early attention to the role of social norms in human conduct (or to the controversy surrounding it), I had always counted myself squarely in the camp of its proponents – those who judged that role to be systematic and consequential. But, as I've stated, my view hadn't developed from any close reading of the pertinent literature but, instead, from a set of informal personal observations associated with some accidents of birth. Once I undertook a full and dispassionate assessment of the literature, I had to admit that those who disagreed with my position had a point. In fact, they had a pair of worthy points. The first went as follows: "Well, look, there are social norms that directly contradict one another, such as the norm for minding your own business and the norm for getting involved. So, no matter which behavior occurred, it could be later explained as normative." This struck me as a troublesome matter for a social norms approach, as I had long since recognized that accounts that could explain everything after the fact were probably too vague or circular to explain anything. The second kind of objection to norms-based explanations was different from the first but equally problematic. In this case, critics pointed out that many times in a society people act counternormatively (e.g., declining to provide assistance to a needy other). Therefore, the critics argued, when people do perform normatively (e.g., agreeing to provide assistance), why should we believe it was because of a societal norm? Wasn't that norm also in place within the culture when they behaved counternormatively? Thus, the critics concluded that we would be better advised to attribute either kind of behavior to the action of other factors.

After immersing myself in the relevant research literature and thinking about the issues for a while, it appeared to me that both sides of the debate were right: norms do have a strong and regular impact on behavior, but the force and form of that impact could only [p. 297 ↓] be clearly established by making certain theoretical advancements. The first involved a crucial conceptual distinction.

Descriptive versus Injunctive Norms

In both everyday parlance and academic usage, the term “norm” has two meanings. One meaning refers to what is typically done (i.e., what is *normal*) in a culture or subculture; my coworkers and I have termed these *descriptive norms*. Another meaning refers to what is typically approved/disapproved within a culture or subculture; we have termed these *injunctive norms*. Despite the common label, the two types of norms come from quite different sources of motivation. Descriptive norms (sometimes called the norms of “is”) motivate by informing individuals of what is likely to be effective and adaptive action in a situation. As such, they provide a decision-making shortcut and information-processing advantage when one is choosing how to behave in a particular setting. By simply registering what most others are doing there and following suit, one can usually choose efficiently and well (Surowiecki, 2004). In contrast to descriptive norms, which specify what is done, injunctive norms specify what ought to be done. They constitute the moral rules of the group, and they motivate action by promising to provide or withhold a form of social acceptance, which is a potent spur for behavior (Williams, 2007). Hence, whereas descriptive norms inform behavior, injunctive norms enjoin it.

Norm Focus

There is substantial evidence that shifting an individual's attention to a specific source of information or motivation will change the individual's responses in ways that are congruent with the features of the now more prominent source (Kallgren and Wood, 1986; Lassiter et al., 2007; Millar and Tesser, 1989; Oyserman and Lee, 2008; Wilson and Gilbert, 2008). In keeping with such evidence, Deaux and Major (1987) concluded that the occurrence of gender-consistent behavior is frequently determined by situational factors that shift attention to the construct of gender, thereby making it more salient. A similar relationship appears to be obtained in the normative arena. That is, norms motivate and direct action primarily when they are activated (i.e., made salient or otherwise focused upon); thus, persons who are dispositionally or temporarily

focused on normative considerations are decidedly more likely to act in norm-consistent ways (Berkowitz, 1972; Kallgren et al., 2000; Miller and Grush, 1986).

An analysis of this sort allows us to retain a belief in the usefulness of normative explanations in the face of the insightful criticisms discussed earlier. That is, it becomes wholly understandable why the dominant norms of a society – that are presumably always in place – may only sometimes predict behavior: they should activate behavior only when *they* have been activated first. Similarly, the simultaneous existence of incompatible social norms is no longer a damaging criticism of normative accounts if we assume that the conflicting norms may coexist within the same society, but that the one that will produce congruent action is the one that is temporarily prominent in consciousness.

Pursuing this last realization further, we can see that it also applies to the distinction between descriptive and injunctive norms. Although it is most frequently the case that what is done and what is approved in a social group are the same, this is often not the case. For instance, even though the majority of people who pass a sidewalk Salvation Army donation kettle might not give a contribution, it is likely that the majority would approve of someone who did. In situations of this kind, with clearly conflicting descriptive and injunctive norms, we would expect that focusing observers on what most people did or on what most people approved would lead to behavior change that is consistent only with whichever has become now the more salient type of norm.

[p. 298 ↓]

Normative Conduct in the Field

From the outset, our intention was to test our theoretical model as it applied to individuals' behavioral decisions regarding the environment. We chose environmental action as our primary dependent measure because it allowed us to test our norm focus model on behavior that we felt was of looming practical importance – although at the time we began the work, in the mid 1970s, that view was not widely shared. More recently, things have changed dramatically. For a variety of reasons (e.g., dwindling supplies of non-renewable energy, worrisome climate change data, concern

for the welfare of future generations, and a general reverence for nature), numerous organizations have urged citizens toward a proenvironmental stance and away from environmentally wasteful or damaging activities. Plainly, environmental action has proven to be a social issue worthy of study. Nonetheless, when we began the work, our motive was not first and foremost to perform good social service; it was to conduct good social science, which required that we undertake proper tests of our model.

To substantiate the need for the theoretical refinements presented in our norm focus model, two questions needed to be answered: (1) Do behavioral patterns confirm our theorized distinction between descriptive and injunctive norms? And (2) is focus a critical mediator of which type of norm guides behavior? Depending upon how these questions are answered, there is also a third question of the practical implications and applications of our theoretical formulation. To attempt to converge upon the answers to these questions, we began with a set of experiments that examined tendencies to litter in public places. Even though littering is not the worst of environmental sins, it had the advantages of involving a discrete act that was easy to measure and was roundly viewed as counter to existing norms.

We thought it was important to conduct our studies in field settings where littering would occur naturally. Although people will litter in laboratory contexts (e.g., Krauss et al., 1978), the external validity of such studies might be questioned. Given the stormy history surrounding the practical utility of normative explanations, we wanted to maximize our external validity in order to offer suggestions for environmental action programs. Thus, we conducted the bulk of our research in field settings to increase our ability to generalize to real-world situations.

The Effects of Focusing on Descriptive Norms

We first turned our attention toward the explication of the effects of focusing on the descriptive norms of a situation. One of the most commonly reported findings from studies of littering is that individuals litter into an already littered environment at a greater rate than they do into an otherwise clean environment. According to our focus

theory, this occurs because individuals are to some degree focused on the descriptive norms present in the situation; that is, they can see the amount of litter already there. Of course, our model is not the only one capable of explaining this data pattern. A social learning theorist might say that the effect is due to subjects' imitation of the behavior of those who have been in the environment before them. Consequently, in order to show the utility of our theoretical refinements, we needed to develop a theoretical test that would predict effects for our model that were different from those predicted from the imitation-based alternative account.

Study 1

In one investigation (Cialdini et al., 1990, Experiment 1) that took place in a hospital parking garage, participants were given the opportunity to litter (a handbill they found on their car windshields) either into a previously clean or a fully littered environment after first witnessing a confederate who either dropped trash into the environment or who simply walked through it. By varying the [p. 299 ↓] state of the environment (clean versus littered), we sought to manipulate the perceived descriptive norm for littering in the situation. By manipulating whether the confederate dropped trash into the environment, we sought to differentially focus participants' attention on the state of the environment and, consequently, to manipulate the salience of the perceived descriptive norm there (i.e., what most people did).

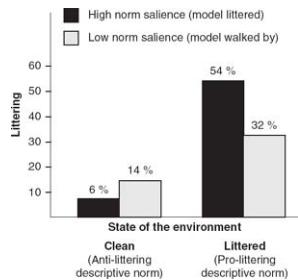
We had three main predictions. First, we expected that participants would be more likely to litter into an already littered environment than into a clean one. Second, we expected that participants who saw the confederate drop trash into a fully littered environment would be most likely to litter there themselves, because they would have had their attention drawn to evidence of a prolittering descriptive norm – that is, to the fact that people typically litter in that setting. Conversely, we anticipated that participants who saw the confederate drop trash into a clean environment would be least likely to litter there, because they would have had their attention drawn to evidence of an antilittering descriptive norm – that is, to the fact that (except for the confederate) people typically do not litter in that setting. This last expectation distinguished our normative account from explanations based on simple modeling processes in that we were making the ironic prediction of decreased littering after participants had witnessed a model litter.

As can be seen in [Figure 41.1](#), the data pattern supported our experimental hypotheses. Overall, there was more littering in the littered environment than in the clean environment. In addition, the most littering occurred when participants saw a model drop trash into a littered environment, and, most tellingly, the least littering occurred when participants saw a model drop trash into a clean environment. Counterintuitive findings of this sort call out for replication. Consequently, we conducted a replication and extension of Study 1 in order to detect the hypothesized decrease in littering when subjects were focused on litter in an otherwise clean environment. This second study (Cialdini et al., 1990, Experiment 2) was also designed to determine if the results from Study 1 were generalizable to other settings and other focus manipulations or whether they were due to some unique characteristics of our previous study.

Study 2

We reasoned that a lone piece of litter would, by its conspicuous nature, draw attention to [p. 300 ↓] the nearly pristine state of the environment. Thus, we expected subjects' littering would decrease when the amount of litter in the environment increased from zero to one piece because the single piece of litter would serve to focus subjects on the antilittering descriptive norm. As the number of pieces of litter in the environment increased beyond one, however, the perceived descriptive norm would change from antilittering to prolittering. As the descriptive norm changed in this fashion, we expected the littering rate would increase. Thus, we made a counterintuitive prediction that could be best described graphically as a check mark-shaped relationship between amount of existing litter in the environment and the likelihood that subjects would litter into it.

Figure 41.1 Percentage of participants littering as a function of descriptive norm salience and the state of the environment in Study 1



To test these hypotheses, we observed the littering tendencies of adult visitors to an amusement park. At one-minute intervals, the first adult to pass a confederate was given a handbill that read "DON'T MISS TONIGHT'S SHOW." Immediately afterward, upon rounding a corner, subjects were unobtrusively observed by a different experimenter as they walked down a path of approximately 55 meters (60 yards) on which we had placed 0, 1, 2, 4, 8, or 16 clearly visible handbills. All other litter had been removed from the walkway.

A visual inspection of [Figure 41.2](#) reveals a pattern of results that is consistent with our predictions. Indeed, a planned comparison using trend weights for a check mark function (-2, -4, -1, 1, 2, 4) was significant. Most notable was the finding that littering decreased (from 18 percent to 10 percent) when one piece of litter was added to a litter-free environment.

Interim Summary

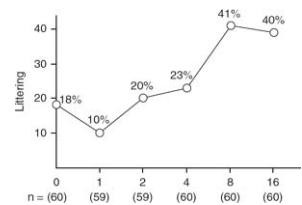
At this juncture, it appeared to us that one factor that motivated our participants' decisions to litter was the descriptive norm of the situation. That is, under control conditions, they littered more in littered environments (where the descriptive norm favored littering) than they did in clean environments (where the descriptive norm opposed littering). More important, when the saliency of these descriptive norms was increased, our participants tended to litter even more in littered environments and even less in clean environments. From a theoretical perspective, then, focusing people on the appropriate descriptive norm regarding littering increased their norm-consistent behavior regardless of whether the norm favored littering or not littering. With this

pattern documented, we turned our focus toward the behavioral influence of injunctive norms.

Effects of Focusing on Injunctive Norms

Recall that, in keeping with the focus theory of normative conduct, our position has been [p. 301 ↓] that norm theorists must be specific about whether they are referring to the descriptive or injunctive norm and which of these, if either, is salient. Up to this point we had only presented supportive evidence in the realm of descriptive norms. A demonstration that focusing on the injunctive norm against littering leads to injunctive norm-consistent behavior (e.g., decreased littering) would be theoretically as well as practically important.

Figure 41.2 Percentage of participants littering as a function of the number of pieces of litter in the environment in Study 2



Study 3

To focus people solely on the injunctive norm against littering in a way that was not amenable to alternative explanations, we relied on the effect of cognitive priming (which, by this date in our theory development, had already been established; see Higgins and Bargh, 1987). That is, one concept (e.g., littering) has a greater probability of being activated when attention is drawn to a related concept (e.g., recycling) compared to when attention is drawn to an unrelated concept (e.g., fine arts). Furthermore, many explanations of priming effects invoke the concept of spreading activation, which posits that similar concepts are linked together in memory within a network of nodes and that activation of one concept results in the spreading of the activation along the network to

other semantically or conceptually related concepts. If norms are stored in a network format, as was suggested by Harvey and Enzle (1981), then varying the relatedness of activated norms to the target antilittering injunctive norm should result in systematic variations in the activation of the target norm. As the relatedness of the activated norm to the antilittering norm increases, the strength of activation of the target norm should also increase and littering rates should decline.

In selecting which other norms to activate, we considered not only the cognitive similarity between the selected norms and the antilittering norm but also the injunctive normativeness (likelihood that violation of it would meet with disapproval from others) of the selected norms. We based our selections on the results of the following scaling procedure. A total of 35 possible norms, including the antilittering norm, were presented to two separate classes of upper division psychology students. The first class rated the 35 possibilities as to their normativeness; the second class rated their conceptual similarity to the antilittering norm. Based on these ratings, we selected four norms that were comparable in perceived normativeness to the antilittering norm; however, one was identical to the antilittering norm (refraining from littering), one was rated close to the antilittering norm (recycling), another was rated moderately close to the antilittering norm (turning out lights), and another was rated far from the antilittering norm (voting). We also selected a control issue that was non-normative (the availability of museums).

Our experimental setting was a community library parking lot. We left the extant litter in place (there was a small amount of litter that was equivalent across all conditions). To manipulate focus on the various norms, we tucked flyers with norm-relevant statements under the driver's side windshield wiper of each car while the patrons were in the library. Upon returning to their cars, subjects found handbills on their windshields with one of the following statements that differed in their similarity/closeness to the antilittering norm: "April is Keep Arizona Beautiful Month. Please Do Not Litter." (identical); "April is Preserve Arizona's Natural Resources Month. Please Recycle." (close); "April is Conserve Arizona's Energy Month. Please Turn Off Unnecessary Lights." (moderately close); "April is Arizona's Voter Awareness Month. Please Remember That Your Vote Counts." (far); and "April is Arizona's Fine Arts Month. Please Visit Your Local Art Museum." (control). We unobtrusively recorded littering of these handbills.

As is clearly evidenced in [Figure 41.3](#), a significant linear trend was obtained, as predicted. This trend indicated that as the conceptual distance between the activated norm or concept and the target antilittering norm increased, littering also increased, supporting our contention that it should be possible to [p. 302 ↓] progressively (and significantly) reduce littering by progressively shifting subjects' focus to the injunctive norm.

Middle Years: Logical Extensions and Practical Implications

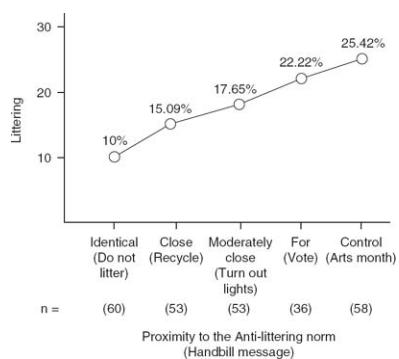
With bolstered confidence in the validity of our model's fundamental elements, we began to think about how those elements could be employed to reduce more general and more damaging environmental problems than litter abatement; for example, how developers of environmentally oriented public service announcements might best structure norm-based messages (Bator and Cialdini, 2000). In addition, we started to test certain logical extensions of the model in settings that would easily lend themselves to beneficial practical applications.

For instance, a conceptually and practically important upshot of our formulation becomes apparent when communicators seek to persuade an audience to behave in accordance with existing norms. As we've stressed, for information campaigns to be successful, their creators must recognize the distinct power of descriptive and injunctive norms and must focus the target audience only on the type of norm that is consistent with the goal. This is far from always the case. We recognized that there was an understandable but misguided tendency of public officials to try to mobilize action against socially disapproved conduct by depicting it as regrettably frequent, thereby inadvertently installing a counterproductive descriptive norm in the minds of their audiences. To understand the logic of the error, consider the following incident that spurred a related experiment.

A graduate student of mine had visited the Petrified Forest National Park in Arizona with his fiancée – a woman he described as the most honest person he'd ever known, someone who had never borrowed a paperclip without returning it. They quickly

encountered a park sign warning visitors against stealing petrified wood, "OUR HERITAGE IS BEING VANDALIZED BY [p. 303 ↓] THE THEFT OF 14 TONS OF WOOD EVERY YEAR." While still reading the sign, he was shocked to hear his fiancée whisper, "We'd better get ours now."

Figure 41.3 Percentage of participants littering a handbill message as a function of its proximity to the injunctive norm against littering in Study 3



In a team meeting where he related the incident, we wondered what could have spurred this wholly law-abiding young woman to want to become a thief and to deplete a national treasure in the process? After some discussion, we thought it had to do with a mistake that park officials made when creating the sign. They tried to alert visitors to the park's theft problem by telling them that many other visitors were thieves. In so doing, they stimulated the behavior they had hoped to suppress by making it appear the norm.

Although it is understandable that park officials would want to instigate corrective action by describing the dismaying size of the problem, our model warned such a message would be far from optimal. That is, *in situations already characterized by high levels of socially censured conduct*, the distinction between descriptive and injunctive norms offers a clear implication: it is a serious error to focus an audience on the descriptive norm (i.e., what is done there); instead, under such conditions, public service messages should focus the audience on the injunctive norm (i.e., what is approved/disapproved there). Thus, according to an informed normative account, the park signage would have been better designed to focus visitors on the social disapproval (rather than the harmful prevalence) of environmental theft.

Studies

Study 4

To examine this hypothesis – that in a situation characterized by unfortunate levels of socially disapproved conduct, a message that focuses recipients on the injunctive norm will be superior to messages that focus recipients on the descriptive norm – we gained permission from Petrified Forest National Park officials to place secretly marked pieces of petrified wood along visitor pathways in three park locations. During five consecutive weekends, at the entrance to each path, we displayed signage that emphasized either descriptive or injunctive norms regarding the theft of petrified wood from the park. The descriptive norm sign stated, “Many past visitors have removed petrified wood from the Park, changing the natural state of the Petrified Forest.” This wording was accompanied by pictures of three visitors taking wood. In contrast, the injunctive norm sign stated, “Please don’t remove the petrified wood from the Park, in order to preserve the natural state of the Petrified Forest.” This wording was accompanied by a picture of a lone visitor stealing a piece of wood, with a red circle-and-bar symbol superimposed over his hand. Our measure of message effectiveness was the percentage of marked pieces of wood stolen over the five-week duration of the study. As predicted, the descriptive norm message resulted in significantly more theft than the injunctive norm message, 7.92 percent versus 1.67 percent (Cialdini, 2003).

Should one conclude from these results that highlighting descriptive norms is always likely to be a counterproductive tactic in environmental information campaigns?

No. In contrast to situations in which environmentally harmful behavior is prevalent, highlighting descriptive norms should be effective for those action domains in which environmentally *beneficial* behavior is prevalent. For example, if the majority of citizens conserve energy at home, campaign developers would be well advised to include such descriptive normative information in their presentations intended to increase residential energy conservation. Of course, if the majority of citizens also approved of such efforts, the campaign developers would be wise to incorporate this injunctive normative information as well.

Study 5a

Thus, the most effective norm-based persuasive approach under these circumstances would be one that enlisted the conjoint influence of descriptive and injunctive norms. [p. 304 ↓] To examine the impact of an information campaign that combined the influence of injunctive and descriptive norms, my colleagues and I created three public service announcements (PSAs) designed to increase recycling, an activity that was truly both performed and approved by the majority of local residents. Each PSA portrayed a scene in which the majority of depicted individuals engaged in recycling, spoke approvingly of it, and spoke disparagingly of a single individual in the scene who failed to recycle. When, in a field test, these PSAs were played on the local TV and radio stations of four Arizona communities, a 25.35 percent net advantage in recycling tonnage was recorded over a pair of control communities not exposed to the PSAs but whose recycling was also measured during the length of the study.

Although a 25 percent recycling advantage is impressive from a practical standpoint, that study did not allow for confident theoretical conclusions about the causes of the advantage. For instance, it was not possible to determine the extent to which our PSAs may have been effective because of their normative elements. After all, it is conceivable that the PSAs had been successful because they included humorous and informational components unrelated to norms, as both humor and novel information have been shown to increase the persuasiveness of a communication when properly employed. In order to assess whether and to what degree descriptive and injunctive norms – separately and in combination – contribute to message effectiveness, mediational evidence was necessary. To that end, we conducted a study in which college students viewed our three recycling PSAs and rated their impact along several relevant dimensions.

Study 5b

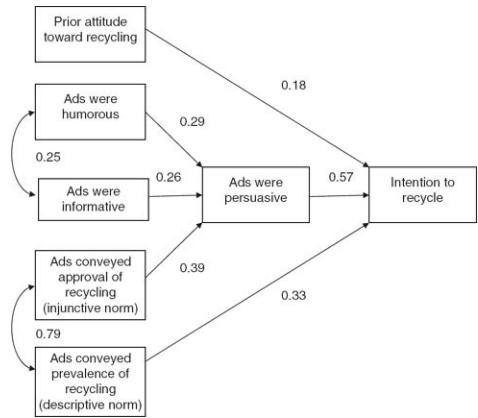
The study (reported in Cialdini, 2003) was designed to determine whether our PSAs had the intended effect of conveying to viewers that recycling was prevalent (descriptive norm) and approved (injunctive norm), whether these perceived norms influenced

viewers' intentions to recycle, and whether the two types of norms operated similarly or differently to affect recycling intentions. This last goal held particular conceptual interest for us. Although considerable research indicates that descriptive and injunctive norms operate independently to affect behavior (Buunk and Bakker, 1995; Okun et al., 2002; Reno et al., 1993), these studies had not examined the mechanisms by which the two types of norms might differ in producing their influence.

Several interesting findings emerged when we examined possible contributors to the effectiveness of the set of PSAs that had been previously successful in stimulating recycling (see [Figure 41.4](#)). First, both normative and non-normative factors influenced the intent to recycle. Of course, the finding that non-normative factors (prior attitude, new information, humor) had causal impact in our data is not incompatible with our theoretical position, as we certainly would not claim that normative factors are the only motivators of human responding.

At the same time, it is encouraging from our theoretical perspective that both injunctive and descriptive normative information significantly influenced recycling intentions. That is, as a result of viewing the ads, the more participants came to believe that recycling was (a) approved and (b) prevalent, the more they planned to recycle in the future. It is noteworthy that, despite a strong correlation ($r = 0.79$) between participants' perceptions of the existing prevalence and approval of recycling, these two sources of motivation had independent effects on recycling intentions and appeared to have these effects via differing psychological mechanisms. Information about social approval/disapproval affected recycling intentions by influencing assessments of communication persuasiveness. Information about relative prevalence, on the other hand, influenced intentions directly, without affecting the perceived persuasiveness of the communication. Such results affirm the theoretical distinction between descriptive and injunctive norms. [p. 305 ↓] Although what most others do and approve are often highly related, they are conceptually, motivationally, and functionally different. Communicators who fail to recognize these distinctions will imperil their persuasive efforts.

Figure 41.4 Intention to recycle after viewing PSAs. The causal paths (arrows) of the figure depict the impact of participants' attitudes and perceptions on their intentions to recycle in Study 5b. All path coefficients (numbers) are significant at $p < 0.05$



Recent Years: Identifying Descriptive–Injunctive Differences and a New Type of Norm

By this stage of model development, one practical implication of our formulation seemed clear: public service communicators should avoid the tendency to send the normatively muddled message that a targeted activity is socially disapproved but widespread. Norm-based persuasive communications are likely to have their best effects when communicators align descriptive and injunctive normative messages to work in tandem rather than in competition with one another. Such a line of attack unites the power of two independent sources of normative motivation and can provide a highly successful approach to social influence.

At the same time, certain issues remained to be clarified if we were to advance theory development and the potential impact of norm-based communications. The most prominent concerned the nature of the psychological mechanisms that underlie descriptive and injunctive norms. The results of Study 5 suggested an intriguing difference. Information about social approval/disapproval affected recycling intentions by influencing assessments of communication persuasiveness. Information about relative prevalence, on the other hand, influenced intentions directly, without affecting

the perceived persuasiveness of the communication. [p. 306 ↓] Why should that be the case? One possibility is that – in contrast to injunctive norms that are based in an understanding of the moral rules of the society – because descriptive norms are based in perceptions of the raw behavior of others, it becomes relatively easy to accommodate to the norm without much cognitive analysis or awareness. Indeed, organisms with little cognitive capacity are able to do so: birds flock, fish school, and social insects swarm. Hence, we might expect that targets of descriptive norm information would be both significantly affected by it and cognitively unaware of its impact. Evidence in this regard comes from a study of perceived motivations for energy conservation that Wesley Schultz and I conducted recently along with our then-students Jessica Nolan, Noah Goldstein, and Vladas Griskevicius.

Studies

Study 6a

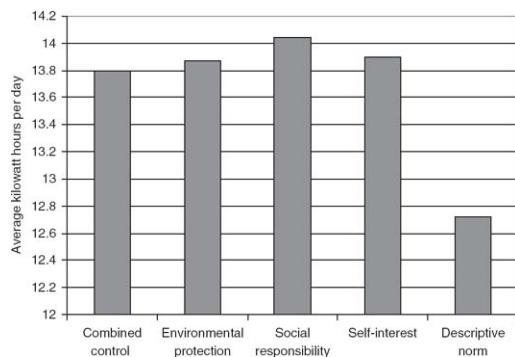
As part of a large-scale survey of residential energy users, we inquired into respondents' views of their reasons for conserving energy at home as well as reports of their actual residential energy-saving activities such as installing energy-efficient appliances and light bulbs, adjusting thermostats, and turning off lights. When respondents were asked to rate the importance to them of several reasons for energy conservation – because it will help save the environments, because it will benefit society, because it will save me money, or because other people are doing it – they rated these motivations in the order just listed, with the actions of others (the descriptive social norm) in a distant last place. However, when we examined the relationship between participants' beliefs in these reasons and their attempts to save energy, we found the reverse: the belief that others were conserving correlated twice as highly with reported energy-saving efforts than did any of the reasons that had been rated as more important personal motivators.

Study 6b

To ensure that our findings weren't the result of the correlational nature of the survey methodology (through a false consensus effect), a follow-up study employed an experimental design. Residents of a mid-size California community received persuasive appeals on door-hangers placed on their doorknobs once a week for four consecutive weeks. The appeals emphasized to residents that energy conservation efforts would (1) help the environment or (2) benefit society or (3) save them money or (4) were common (normative) in their neighborhood. We also included a no-contact control group and a control group that received door-hangers that simply urged energy conservation without providing a specific reason; these controls did not differ from one another in impact on usage. Interviews with participants revealed that those who received the normative appeals rated them as least likely to motivate their conservation behavior and significantly less likely than any of the other appeals. Yet, when we examined actual energy usage (by recording participants' electricity meter readings), the normative appeal proved most helpful, resulting in significantly more conservation than any of the other appeals (Nolan et al., 2008; see [Figure 41.5](#)). This illustrates a more specific psychological point than the one articulated masterfully by Nisbett and Wilson (1977) that, in general, people are poor at recognizing why they behave as they do. It asserts that they will be particularly clueless when identifying the similar actions of others as causal antecedents. The upshot of these studies is plain. When it comes to estimating the causes of their conduct, people seem especially blind to the large relative role of others' similar behavior. They don't just fail to get this relative role right; they tend to get it precisely wrong.

The tendency to underestimate the power of descriptive normative information doesn't just apply to the recipients of the information. It applies to potential deliverers of the information as well. This suggests a simple [p. 307 ↓] way to increase conservation activity – by trumpeting the true levels of conservation that are going unrecognized. However, proenvironmental program developers and communicators frequently fail to do so, relying instead on other forms of motivation than normative information (Nolan et al., 2008).

Figure 41.5 Average daily energy consumption as a function of message, controlling for baseline energy usage in Study 6b



Study 7

To investigate range of this error of omission, we examined resource conservation choices in upscale hotel rooms, where guests often encounter a card asking them to reuse their towels. As anyone who travels frequently knows this card may urge the action in various ways. Sometimes it requests compliance for the sake of the environment; sometimes it does so for the sake of future generations; and sometimes it exhorts guests to cooperate with the hotel in order to save resources. What the card *never* says, however, is that the great majority (up to 75 percent) of guests do reuse their towels when given the opportunity, even though this percentage is accurate according to data from the Project Planet Corporation that manufactures the cards. We suspected that the omission was leading to inferior levels of compliance.

To test our suspicion, with the collaboration of the management of an upscale hotel in the Phoenix area, we put one of four different cards in its guestrooms. One of the cards stated "HELP SAVE THE ENVIRONMENT," which was followed by information stressing respect for nature. A different card stated "HELP SAVE RESOURCES FOR FUTURE GENERATIONS," which was followed by information stressing the importance of saving energy for the future. A third type of card stated "PARTNER WITH US TO HELP SAVE THE ENVIRONMENT," which was followed by information urging guests to cooperate with the hotel in preserving the environment. A final type of card stated "JOIN

YOUR FELLOW CITIZENS IN HELPING TO SAVE THE ENVIRONMENT," which was followed by information that 75 percent of hotel guests do reuse their towels at some point during their stay when asked to do so. The outcome? Compared with the first three messages, the final (descriptive social norm) message increased towel reuse by an average of 28.4 percent (Goldstein et al., 2007).

[p. 308 ↓]

We found it instructive that the descriptive normative message was (a) significantly more successful than any of the more traditionally employed appeals, (b) costless to the hotel, (c) entirely honest, yet (d) never employed in any hotel we had ever visited. We think that there is a good reason why this would be the case. Recall that Studies 6a and 6b provided evidence that people tend to dismiss the impact that descriptive norms can have on their own actions. Therefore, it may well be that when hotel energy conservation program creators and communicators consulted their own phenomenologies to determine what would likely be a strong appeal (i.e., what would likely work on them), it didn't occur to them that descriptive norms would be effective. As one form of corrective to this misjudgment, even before submitting a report of our hotel findings to an academic journal, we have published them in an outlet read by hotel managers and executives (Goldstein et al., 2007).

Whose Descriptive Norms Do Individuals Follow? The Emergence of Provincial Norms

To this point, we have considered when and how individuals adhere to the perceived norms of others, but one central question remains to be addressed: Whose norms are individuals most likely to follow? Social psychologists have long recognized that when making decisions under uncertainty, people tend to follow the norms of others who seem similar to them (Baron et al., 1996; Festinger, 1954; Sherif, 1936). By far the most frequently studied form of similarity has involved common membership in some personally meaningful group or social category such as race, gender, nationality, or

ethnicity (Brewer, 2003; Hogg, 2003; Terry and Hogg, 2000). However, another line of research suggests that similarities of a personally meaningless variety – for example, fingerprint type or birthday – are powerful stimulators of compliance with another's requests (Burger et al., 2004). This latter set of findings made us think that the norms of others who are similar to an observer on a *social identity-irrelevant* dimension might still lead the observer to elevated levels of conformity with those norms. What might that social identity-irrelevant yet powerful dimension of similarity be? We suspected that individuals would follow the behavioral norms of those who share or have shared the environment in which a behavior is to be undertaken.

This suspicion made functional sense to us. When deciding which action to take in a specific environment, it would be adaptive to follow the lead of others who had taken action in that specific environment, even if those individuals were not members of one's meaningful social categories. We labeled the norms of such individuals *provincial norms* to denote simultaneously their local and socially unsophisticated nature. Despite their geographically and interpersonally narrow character, we nonetheless expected provincial norms to be potent directors of conduct because they can be seen to provide highly diagnostic information about wise behavioral choices in the situation at hand.

Study 8

Recall that the descriptive norm used in our hotel study informed guests that similar others – that is, the majority of other guests who had previously stayed at the hotel – had reused their towels at least once during their respective stays. We decided to explore the power of provincial norms by conducting another hotel study in which some occupants saw an appeal communicating the descriptive norm specifically for the guests who had previously stayed in the occupants' hotel rooms (i.e., the provincial norm for reusing towels). Thus, in addition to the standard environmental appeal of our earlier hotel study and the descriptive norm appeal used in that study, participants in a third condition read that “75% of the guests who stayed in this room [room #xxx] participated in our new resource savings program by using their towels more than once.”

[p. 309 ↓]

Consistent with our first hotel study (Study 7), the descriptive norm condition using the hotel's previous guests as the reference group yielded a higher towel reuse rate than did the standard environmental appeal - 18.3 percent higher in this instance. More interesting, however, was the finding that the provincial norm condition using the room's previous occupants as the standard produced an even higher towel reuse rate (32.5 percent) relative to the standard environmental appeal (Goldstein et al., 2008).

It may seem illogical that guests would be more likely to follow the norms of those who previously stayed in their specific room than those who stayed in other rooms throughout the hotel. That is, from a purely rational standpoint, one should not view the previous occupants of one's room in a positive light. After all, these are the same individuals who have reduced the quality of the room and its amenities for the current occupants. In addition, there is no ready reason to believe that the behaviors of those previously occupying one's room are more valid than the behaviors of those previously occupying the room next door. Indeed, because the number of guests previously staying at the hotel constitutes a larger sample than the number of guests previously staying in any one room, the former type of norm should be more instructive than the latter. Yet, the reverse was the case.

Why? It is generally adaptive for one to follow the behavioral norms associated with the particular environment, situation, or circumstances that most closely match one's own environment, situation, or circumstances. Thus, individuals may develop this general tendency into a mental shortcut, which, like other mental shortcuts, can sometimes lead to judgments, decisions, and behaviors not entirely based on logical analysis (Ariely, 2008; Thaler and Sunstein, 2008; Tversky and Kahneman, 1974). The findings from our research suggest that communicators, policymakers, and managers incorporating a descriptive normative component into their persuasive appeals or information campaigns should ensure that the group on which the norms are based is not just comparable to the intended audience in socially meaningful ways (e.g., age, gender, or ethnicity) but is also comparable in situational and circumstantial ways.

Conclusion

The focus theory of normative conduct was launched as a way to make sense of the mixed support at the time for the role of social norms in human behavior. The claim that the concept, as traditionally conceived, possessed great explanatory power had strong proponents and equally strong opponents. From the perspective of the research my colleagues and I have since conducted on the question, it would appear that both camps were right. Norms clearly do have a considerable impact on behavior, but the force and form of that impact can only be usefully understood through certain conceptual elements that had not been traditionally or rigorously applied. That is, to predict properly the likelihood of norm-consistent action requires, first, that one specify the type of norm – descriptive or injunctive – said to be operating. Second, one must take into account the various conditions that would incline individuals to focus attention onto or away from the norm, as it is only when either type of norm is salient that it is likely to direct behavior forcefully. Early research confirmed the importance of both of these conceptual elements in generating norm-consistent conduct.

Because of a desire to advance the theory in both conceptual and practical directions, subsequent research was designed to test certain logical extensions of the model in settings that would easily lend themselves to beneficial practical applications. For example, one implication of the conceptual distinction between injunctive and descriptive norms is that, in a situation characterized by unfortunate levels of socially disapproved conduct, it is a serious error to focus an [p. 310 ↓] audience on the descriptive norm (i.e., what is done there); instead, under such conditions, messages should focus the audience on the injunctive norm (i.e., what is approved/disapproved there). A study of signage at the Petrified Forest National Park supported this inference in demonstrating a reduction in theft of petrified wood from an injunctive normative message but an increase in theft from a descriptive normative message. Later work tested a theory-based set of PSA messages and found that descriptive and injunctive components of the messages operated separately and via different psychological processes to increase household recycling intentions and behavior. More recent experiments of signage urging towel reuse in hotel rooms revealed (a) the superiority of norms-based messages to traditional messages and (b) evidence for the special power

of provincial norms – descriptive norms of others who have shared precisely the same locale as an observer.

There are several potentially fruitful directions that future research can take. One such direction is the continued exploration of the differences between descriptive and injunctive norms. For instance, because injunctive norms are said to require more cognitive work to operate, they should be less effective than descriptive norms when individuals have recently expended considerable cognitive effort (i.e., are cognitively depleted). Some as yet unpublished evidence supports this expectation: information regarding what most others approved (the injunctive norm) spurred norm-consistent behavior when presented before a strenuous, cognitively depleting activity but not after the activity; in contrast, information regarding what most others did (the descriptive norm) increased norm-consistent behavior whenever it was presented (Jacobson et al., 2009). Another potentially worthwhile program of work would investigate the question of why the norms of similar others lead to norm-consistent conduct. We have suggested that for provincial norms, which engage no sense of meaningful social identification with the similar others, it is the diagnostic/instructional value of the others' actions that motivates conformity. It is worth asking whether this diagnostic/instructional factor (i.e., the prediction of adaptive outcomes) also leads to conformity with the actions of others' with whom one does share a meaningful social identity. That is, would conformity with the norms of such others (e.g., fellow students at a specific university) occur not because of a sense of identification with those others but because of a sense that their norms are indicative of successful outcomes for anyone like them? Finally, Higgins (pers. comm., May 7, 2009) suggested that descriptive norms may be more associated with the psychological state of assessment, wherein people withhold action until they are confident of the correctness of their views (Higgins et al., 2003), whereas injunctive norms may be more associated with the motivational state of prevention, wherein people are more concerned with avoiding inappropriate conduct (Higgins, 1998). Thus, experimentally inducing either assessment or prevention states may make descriptive or injunctive norms, respectively, more influential. Researchers who wish to pursue these or other implications of the focus theory of normative conduct can be encouraged in this regard because, after all, many others are likely to be doing so.

Robert B.Cialdini

References

- Ariely, D. (2008) *Predictably Irrational*. New York: HarperCollins.
- Baron, R.S., Vandello and Brunsman, B. The forgotten variable in conformity research: Impact of task importance on social influence. *Journal of Personality and Social Psychology*, 71, 915–927. (1996) <http://dx.doi.org/10.1037/0022-3514.71.5.915>
- Bator, R.J. and Cialdini, R.B. The application of persuasion theory to the development of effective pro-environmental public service announcements. *Journal of Social Issues*, 56, 527–541. (2000) <http://dx.doi.org/10.1111/0022-4537.00182>
- Bator, R.J. and Cialdini, R.B. The nature of consistency motivation: Consistency, aconsistency, and anticonsistency in a dissonance paradigm. *Social Influence*, 1, 208–233. (2006) <http://dx.doi.org/10.1080/15534510600823638>
- Berkowitz, L. (1972) Social norms, feelings, and other factors affecting helping and altruism. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, 6, 63–108. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960025-8>
- Brewer, M. (2003) Optimal distinctiveness, social identity, and the self. In M.R. Leary, ed. and J.P. Tangney (eds), *Handbook of Self and Identity*, pp. 480–491. New York: Guilford Press.
- Burger, J.M., Messian, N., Patel, S., del Prado, A. and Anderson, C. What a coincidence! The effects of incidental similarity on compliance. *Personality and Social Psychology Bulletin*, 30, 35–43. (2004). <http://dx.doi.org/10.1177/014616720325838>
- Buunk, B.P. and Bakker, A.B. Extradyadic sex: The role of descriptive and injunctive norms. *Journal of Sex Research*, 32, 313–318. (1995) <http://dx.doi.org/10.1080/00224499509551804>

Cialdini, R.B. Crafting normative messages to protect the environment . Current Directions in Psychological Science , 12, 105–109. (2003) <http://dx.doi.org/10.1111/1467-8721.01242>

Cialdini, R.B., Reno, R.R. and Kallgren, C.A. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places . Journal of Personality and Social Psychology , 58, 1015–1026. (1990) <http://dx.doi.org/10.1037/0022-3514.58.6.1015>

Cialdini, R.B., Trost, M.R. and Newsom, J.T. Preference for consistency: The development of a valid measure and the discovery of surprising behavioral implications . Journal of Personality and Social Psychology , 69, 318–328. (1995) <http://dx.doi.org/10.1037/0022-3514.69.2.318>

Cialdini, R.B., Wosinska, W., Barrett, D.W., Butner, J. and Gornik-Durose, M. Compliance with a request in two cultures: The differential influence of social proof and commitment/consistency on collectivists and individualists . Personality and Social Psychology Bulletin , 25, 1242–1253. (1999) <http://dx.doi.org/10.1177/0146167299258006>

Darley, J.M. and Latané, B. (1970) Norms and normative behavior: Field studies of social interdependence . In J. Macaulay, ed. and L. Berkowitz (eds), Altruism and Helping Behavior , pp. 83–102. New York: Academic Press.

Deaux, K. and Major, B. Putting gender into context: An interactive model of gender-related behavior . Psychological Review , 94, 369–389. (1987) <http://dx.doi.org/10.1037/0033-295X.94.3.369>

Festinger, L. A theory of social comparison processes . Human Relations , 7, 117–140. (1954) <http://dx.doi.org/10.1177/001872675400700202>

Fishbein and Ajzen (1975) Belief, Attitude, Intention, and Behavior . Reading, MA: Addison-Wesley.

Goldstein, N.J., Cialdini, R.B. and Griskevicius, V. Rooms for improvement . Cornell Hotel and Restaurant Administration Quarterly , 48, 145–150. (2007) <http://dx.doi.org/10.1177/0010880407299542>

Goldstein, N.J., Cialdini, R.B. and Griskevicius, V. A room with a viewpoint: Using normative appeals to motivate environmental conservation in a hotel setting . Journal of Consumer Research , 35, 472–482. (2008) <http://dx.doi.org/10.1086/586910>

Harvey, M.D. and Enzle, M.E. A cognitive model of social norms for understanding the transgression-helping effect . Journal of Personality and Social Psychology , 41, 866–875. (1981) <http://dx.doi.org/10.1037/0022-3514.41.5.866>

Higgins, E.T. (1998) Promotion and prevention: Regulatory focus as a motivational principle . In: M.P. Zanna (ed.), Advances in Experimental Social Psychology , 30, 1–46. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960381-0>

Higgins, E.T. and Bargh, J.E. Social cognition and social perception . Annual Review of Psychology , 38, 369–425. (1987) <http://dx.doi.org/10.1146/annurev.ps.38.020187.002101>

Higgins, E.T., Kruglanski, A.W. and PierroA. (2003) Regulatory mode: Locomotion and assessment as distinct orientations . In M.P. Zanna (ed.), Advances in Experimental Social Psychology , 35, 293–344. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2803%2901005-0>

Hogg, M.A. (2003) Social identity . In M.R. Leary, ed. and J.P. Tangney (eds), Handbook of Self and Identity , pp. 462–479. New York: Guilford Press.

Jacobson, R.P., Mortensen, C.R. and Cialdini, R.B. (2009) The role of self-regulation in normative influence . Poster presented at the Meeting of the Society of Personality and Social Psychology, February, Tampa, FL.

Kallgren, C.A., Reno, R.R. and Cialdini, R.B. A focus theory of normative conduct: When norms do and do not affect behavior . Personality and Social Psychology Bulletin , 26, 1002–1012. (2000) <http://dx.doi.org/10.1177/0146167200261009>

Kallgren, C.A. and Wood, W.W. Access to attitude-relevant information in memory as a determinant of attitude-behavior consistency . *Journal of Experimental Social Psychology* , 22, 328–338. (1986) <http://dx.doi.org/10.1016/0022-1031%2886%2990018-1>

Krauss, R.M., Freedman, J.L. and Whitcup, M. Field and laboratory studies of littering . *Journal of Experimental Social Psychology* , 14, 109–122. (1978) <http://dx.doi.org/10.1016/0022-1031%2878%2990064-1>

Krebs, D.L. Altruism: An examination of the concept and a review of the literature . *Psychological Bulletin* , 73, 258–302. (1970) <http://dx.doi.org/10.1037/h0028987>

Krebs, D.L. and Miller, D.T. (1985) Altruism and aggression . In G. Lindzey, ed. and E. Aronson (eds), *The Handbook of Social Psychology*, Vol. 2 , 3rd Edition, pp. 1–71. New York: Random House.

Lassiter, G.D., Diamond, S.S., Schmidt, H.C. and Elek, J.K. Evaluating videotaped confessions . *Psychological Science* , 18, 224–226. (2007) <http://dx.doi.org/10.1111/j.1467-9280.2007.01879.x>

Marini, M.M. Age and sequencing norms in the transition to adulthood . *Social Forces* , 63, 229–244. (1984)

Millar, M.G. and Tesser, A. The effects of affective-cognitive consistency and thought on the attitude-behavior relation . *Journal of Experimental Social Psychology* , 25, 189–202. (1989) <http://dx.doi.org/10.1016/0022-1031%2889%2990012-7>

Miller, L.E. and Grush, J.E. Individual differences in attitudinal versus normative determination of behavior . *Journal of Experimental and Social Psychology* , 22, 190–202. (1986) <http://dx.doi.org/10.1016/0022-1031%2886%2990023-5>

Nisbett, R.E. and Wilson, T.D. Telling more than we can know: Verbal report on mental processes . *Psychological Review* , 84, 231–259. (1977) <http://dx.doi.org/10.1037/0033-295X.84.3.231>

Nolan, J.M., Schultz, P.W., Cialdini, R.B., Goldstein, N.J. and Griskevicius, V. Normative social influence is underdetected . Personality and Social Psychology Bulletin , 34, 913–923. (2008) <http://dx.doi.org/10.1177/0146167208316691>

Okun, M.A., Karoly, P. and Lutz, R. Clarifying the contribution of subjective norm to predicting leisure-time exercise . American Journal of Health Behavior , 26, 296–305. (2002) <http://dx.doi.org/10.5993/AJHB.26.4.6>

Oyserman, D. and Lee, S.W.S. Does culture influence what and how we think? Psychological Bulletin , 134, 311–342. (2008) <http://dx.doi.org/10.1037/0033-2909.134.2.311>

Pepitone, A. Toward a normative and comparative biocultural social psychology . Journal of Personality and Social Psychology , 34, 641–653. (1976) <http://dx.doi.org/10.1037/0022-3514.34.4.641>

Petrova, P.K., Cialdini, R.B. and Sills, S.J. Personal consistency and compliance across cultures . Journal of Experimental Social Psychology , 43, 104–111. (2007) <http://dx.doi.org/10.1016/j.jesp.2005.04.002>

Reno, R.R., Cialdini, R.B. and Kallgren, C.A. The transsituational influence of social norms . Journal of Personality and Social Psychology , 64, 104–112. (1993) <http://dx.doi.org/10.1037/0022-3514.64.1.104>

Sherif, M. (1936) The Psychology of Social Norms . New York: Harper.

Surowiecki, J. (2004) The Wisdom of Crowds . New York: Doubleday.

Terry, D.J. and HoggM.A. (2000) Attitudes, Behavior, and Social Context: The Role of Norms and Group Membership . New Jersey: Lawrence Erlbaum Associates.

Thaler, R.H. and Sunstein, C.R. (2008) Nudge . New Haven, CT: Yale University Press.

Triandis, H.C. (1977) Interpersonal Behavior . Monterey, CA: Brooks/Cole.

Tversky, A. and Kahneman, D. Judgment under uncertainty: Heuristics and biases .
Science , 185, 1123–1131. (1974) <http://dx.doi.org/10.1126/science.185.4157.1124>

Williams, K.D. Ostracism . Annual Review of Psychology , 58, 425–452. (2007) <http://dx.doi.org/10.1146/annurev.psych.58.110405.085641>

Wilson, T.D. and Gilbert, D.T. Affective forecasting: Knowing what to want .
Current Directions in Psychological Science , 14, 131–134. (2008) <http://dx.doi.org/10.1037/0022-3514.78.5.821>

<http://dx.doi.org/10.4135/9781446249222.n41>

Handbook of Theories of Social Psychology

System Justification Theory

Contributors: John T. Jost & Jojanekke van der Toorn

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "System Justification Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n42>

Print pages: 313-344

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n42>

[p. 313 ↓]

Chapter 42: System Justification Theory

Abstract

According to system justification theory, people are motivated (to varying degrees depending upon situational and dispositional factors) to defend, bolster, and justify prevailing social, economic, and political arrangements (i.e., the status quo). System justification motivation is theorized to manifest itself in a number of different ways (e.g., in terms of stereotyping, ideology, attribution), to occur implicitly (i.e., nonconsciously) as well as explicitly, and to serve underlying epistemic, existential, and relational needs. In this chapter, we trace the historical and intellectual origins of the theory, beginning with a personal narrative of its conceptual and empirical development. We recount major influences and theoretical precursors in philosophy, social theory, and experimental social psychology. We summarize the basic postulates of system justification theory in its current state of development, highlight some illustrative evidence in support of the theory, and discuss a few of its practical consequences.

Introduction

Most individuals participate in an astonishing number and diversity of relationships, groups, networks, and social systems. Even in solitude, our thoughts, feelings, and actions reflect social norms, expectations, and the ties that bind (Hardin and Higgins, 1996; Moscovici, 1988; Sherif, 1936). But what do we do when these relationships carry within them elements of inequality, exploitation, and injustice? Resistance – if not out-and-out rebellion – would seem to be the most obvious or appropriate response to such situations (Gurr, 1970; Hirschman, 1970; Klandermans, 1997; Reicher, 2004; Turner, 2006), but its occurrence is rarer than most would expect. The infrequency of protest, collective action, and other convincing attempts to reshape social systems

to make them more congenial to group interests, including majority group interests, was – throughout the twentieth century – a much-studied, albeit not well understood, phenomenon in philosophy and the social sciences. As we will see in this chapter, authors exemplifying a multiplicity of theoretical traditions have concluded that people, including members of disadvantaged groups, frequently acquiesce in the social order and, in so doing, violate their own [p. 314 ↓] objectively defined social interests (Fromm, 1941; Gramsci, 1971; Hochschild, 1981; Jackman, 1994; Kluegel and Smith, 1986; Lane, 1959; MacKinnon, 1989; Mason, 1971; Moore, 1978; Runciman, 1969; Tyler and McGraw, 1986; Zinn, 1968).

However, the acquiescence of those who are disadvantaged by the status quo was not adequately explained or connected to a comprehensive analysis of thought and (in)action in the social and political sphere more generally. More often than not, it was simply attributed to the passive acceptance of “dominant ideology” (e.g., Abercrombie et al., 1990). Almost without exception, scholars failed to consider the possibility that most individuals – and not just those at “the top” – have a psychological interest (or motivation) to uphold the legitimacy of the social system. This is precisely what system justification theory proposes: to varying degrees (based upon dispositional and situational factors), people are motivated (whether consciously or unconsciously) to defend, bolster, and justify aspects of existing social, economic, and political arrangements (Jost and Banaji, 1994; Jost and Hunyady, 2002, 2005; Jost et al., 2004a).

System justification theory has, from its inception, represented a self-conscious attempt to explain why people so frequently adapt themselves to the societal status quo, rather than pushing for change and social betterment, as so many other theories in social science would portend. System justification theory seeks to integrate insights garnered from different philosophical and scientific perspectives on “false consciousness” and political acquiescence and therefore to function as a kind of “umbrella theory” (Jost and Banaji, 1994; Jost and Hunyady, 2002, 2005; Jost et al., 2001). In this chapter, we begin with a fairly personal narrative – consistent with the editorial objectives of this book – of how system justification theory came to be. Second, we review the historical and intellectual origins of the theory, focusing on major influences and theoretical precursors. Third, we summarize the basic tenets (or postulates) of system justification

theory in its current state of development. Fourth and finally, we consider a few of the theory's practical implications, providing a take-home message of sorts.

Personal Narrative of the Theory's Development

System justification theory began life as a one-page critical reaction paper and, still in its infancy, grew into a term paper entitled "Salvaging exploitation theories of prejudice: Stereotypes as social justification," submitted by first-year doctoral student John Jost in the spring semester of 1991 for Professor Mahzarin Banaji's seminar on stereotyping and prejudice at Yale University. The term paper opened with two quotations, one from Karl Marx and the other from Gordon Allport:

The class which has the means of material production at its disposal, has control at the same time over the means of mental production ...

(Marx and Engels, 1846: 64)

[T]he rationalizing and justifying function of a stereotype exceeds its function as a reflector of group attributes.

(Allport, 1954: 196)

The basic argument of the paper was that

- (a) specific stereotype contents arise because they explain and justify the status quo, especially the exploitation of certain groups of people,
- (b) they are initially promulgated by those members of society who stand to gain the most advantage by preserving the exploitative system, and
- (c) they are eventually spread by virtually all members of society, since stereotypes serve the ideological function of explaining social reality in a way which makes it seem natural and just.

From the start, system justification theory represented an attempt to synthesize and unify two distinct theoretical traditions – one coming from philosophy and social theory in the Marxian tradition, which Jost had studied previously, and the other coming from Lewin, Allport, Tajfel, and their scientific heirs in experimental social psychology.

[p. 315 ↓]

The ideas presented in the paper were enthusiastically received by Professor Banaji and led to several lively discussions during lab meetings, but it was not until the *British Journal of Social Psychology* announced plans for a special issue on the structure and functions of stereotyping that Jost and Banaji resolved to write the first article outlining the basic tenets of a system justification approach to stereotyping and prejudice. The earliest articulations were either ambivalent or agnostic about the notion that members of disadvantaged groups were *motivated* to hold system-justifying stereotypes and ideologies. Rather, processes of persuasion and social learning were largely assumed to account for the apparent acceptance of the status quo on the part of the disadvantaged. Jost and Banaji (1994) ultimately opted for “system justification” over “system rationalization,” because the former term seemed less pejorative.

However, the most distinctive aspect of system justification theory, which was not explicitly stated by any of the theory's many influential predecessors (see [Figure 42.1](#)), was the possibility that even members of disadvantaged groups would (for psychological and ideological reasons) *want* to believe that the existing social system is good, fair, legitimate, and right. This motivational assumption, which to some extent drew on the philosophical concept of self-deception, was suggested (with some hesitation) by Jost and Banaji (1994), but it was not directly investigated until several years later (e.g., Jost et al., 2010; Kay et al., 2009). Although system justification theory focused initially on stereotyping, prejudice, and intergroup relations, it was later expanded to account for other types of social judgments, including attributions, explanations, and rationalizations for social events; perceptions of fairness, legitimacy, deservingness, and entitlement (concerning the self and others); specific social and political attitudes; and, ultimately, full-fledged ideological belief systems (Jost et al., 2003b, 2004a, 2004b).

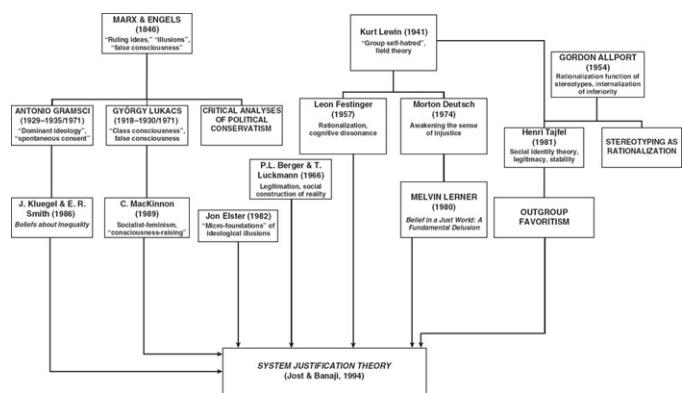
Several members of the Yale faculty (especially Mahzarin Banaji, William J. McGuire, Leonard Doob, and Robert Abelson) encouraged and inspired, either directly or indirectly, the development of system justification theory in its still-fledgling state. Postdoctoral experiences further deepened and broadened Jost's interests in the nature of human attachment to the status quo and the causes of resistance to change. A collaboration with Arie Kruglanski led ultimately to a detailed analysis of political conservatism and its underlying social, cognitive, and motivational underpinnings (Jost et al., 2003b), which provided an empirical basis for Jost and Hunyady (2005) to propose that system-justifying beliefs are appealing in part because they help people to reduce uncertainty and threat. Another postdoctoral collaboration with Brenda Major resulted in an interdisciplinary conference on "The Psychology of Legitimacy" that would eventually take place at Stanford University, where Jost began an assistant professorship in 1997 (see Jost and Major, 2001).

While at Stanford, Jost benefited greatly from interactions with a number of colleagues, but the individual who most advanced the theoretical and empirical development of system justification theory was a doctoral student named Aaron Kay. Their first collaboration, which derived from McGuire and McGuire's (1991) proposal that people engage in "sour grapes" and "sweet lemons" rationalizations, demonstrated that citizens' political preferences were affected by the anticipated status quo; that is, by (manipulated) expectations about which candidate was more likely to win (Kay et al., 2002). Subsequent collaborations addressed the system-justifying potential of complementary stereotypes, such as "poor but happy" and "poor but honest" stereotypes (Jost and Kay, 2005; Kay and Jost, 2003; Kay et al., 2005). Later, as a faculty member at the University of Waterloo, Kay (now at Duke) spearheaded a productive laboratory investigating the aversion to randomness as a motivational antecedent of system-justifying belief systems, including religious ideologies (Kay et al., 2008, 2009).

In 2003 Jost moved to New York University (NYU), where several faculty members were [p. 316 ↓] [p. 317 ↓] focused on the possibility that motivation (or goal systems) can operate implicitly, that is, outside of conscious awareness (Bargh et al., 2001). This inspired Jost and his students to adapt experimental paradigms from social cognition to provide direct support for the notion that system justification follows principles of goal pursuit (Jost et al., 2008b, 2010; Ledgerwood et al., in press). Specifically, we

postulated the existence of an abstract system-justifying goal or motive that can operate both consciously and unconsciously and that leads people to see the societal status quo as relatively legitimate, fair, desirable, and just – that is, as *better* than if it were not the status quo (see also Kay et al., 2009).

Figure 42.1 System justification theory (Jost and Banaji, 1994)



A few other additions to system justification theory are noteworthy. NYU colleague Tom Tyler helped to flesh out (a) the relationship between system justification and other determinants of system legitimacy (e.g., Tyler and Jost, 2007), including procedural fairness and feelings of psychological dependence (van der Toorn et al., 2011), and (b) affective and the relationship between behavioral consequences of system justification, such as the reduction of moral outrage and the withdrawal of support for social change (Wakslak et al., 2007). Furthermore, Jaime Napier (now at Yale) expanded the notion that system justification serves a palliative function to explain why political conservatives are *happier* than liberals (Napier and Jost, 2008). Finally, Jost et al. (2008a) incorporated insights from Hardin and Higgins' (1996) shared reality theory to propose that system justification serves *relational* in addition to epistemic and existential motives (see also Kaiser et al., 2006).

Historical Origins and Intellectual Precursors of System Justification Theory

There are certain historical and intellectual influences – some within the disciplinary boundaries of social psychology and some without – that played truly major roles in stimulating our thinking about how and why people provide ideological support for the status quo, even when it conflicts with personal or group interests (see also Jost et al., 2004b: 264–269; Jost and Hunyady, 2002: 114–118). These influences are summarized visually (in the form of a hypothetical family tree) in [Figure 42.1](#). The fact that these thinkers have probably never been grouped together before speaks rather strongly in favor of Jorge Luis Borges' (1964) observation that “every writer creates his own precursors.” It may well be that the *only* common denominator (other than “social science” in general) to be found in the writings of Marx, Engels, Gramsci, Lukács, Lewin, Allport, Berger and Luckmann, Tajfel, Lerner, Elster, MacKinnon, and the others included in [Figure 42.1](#), is their putative connection to the theoretical construct we have dubbed *system justification*. In fact, we discovered some of these thinkers only *after* the original assumptions of system justification theory had already been laid out; their insights, in other words, were folded into the theory as it developed conceptually as well as empirically.

Approaches to Dominant Ideology and False Consciousness

The concept of system justification is based loosely on the concept of *false consciousness*, which is rooted in the early (humanistic, sociological) work of Karl Marx, especially *The German Ideology* and other works of the 1840s and 1850s (see also Fromm, 1965). In essence, Marx and Engels (1846) argued that ideas favoring dominant groups in society prevail because these groups control the cultural and institutional means by which ideas are spread. The net result, they claimed, is that through the ideological machinations of elites (including philosophers, such as the “young Hegelians”), “men and their circumstances appear upside-down as in a *camera* obscura.”

obscura" (1846: 47); social and political realities, in other words, are systematically [p. 318 ↓] inverted and therefore distorted. But it was Engels who actually used the term "false consciousness" in a letter written several years after Marx's death: "Ideology is a process carried out by the so-called thinker with consciousness, but with a false consciousness. The real driving forces that move him remain unknown to him; otherwise it would not be an ideological process" (e.g., see Runciman, 1969).

Marx believed that the working classes would eventually see through the ideological illusions and strive to overthrow the capitalist system. From the perspective of rational self-interest, Marx emphatically assumed that the poor had "nothing to lose but their chains." However, his expectation that the oppressed would recognize and take action against the sources of their oppression may have been overly optimistic (or, from the perspective of the ruling class, pessimistic), considering the various social and psychological obstacles to social change that exist, including denial, rationalization, and other system justification tendencies (see also Crosby et al., 1989; Jost and Banaji, 1994; Major, 1994).

To explain why revolutions against capitalism (and other arguably exploitative systems) did not occur in heavily industrialized nations, later theorists, most notably Gramsci (1971), Lukács (1971/1989), Fromm (1962), Marcuse (1964), Runciman (1969), Cunningham (1987), and MacKinnon (1989), further developed the analysis of dominant ideology, cultural hegemony, and false consciousness. Each of these constructs, but especially the last one, anticipated the subject matter explored by system justification theorists. However, the concept of system justification was intended to ground these sociological constructs in psychological science and to capture the process rather than simply the *outcome* (or product) of ideological activity. In [Table 42.1](#) we have summarized the major contributions to the system justification perspective made by various theorists with respect to the study of dominant ideology and false consciousness.

Antonio Gramsci, the Marxian social theorist who was imprisoned unto death in Mussolini's fascist regime in Italy, took seriously Marx's notion that a "popular conviction often has the same energy as a material force." More specifically, Gramsci (1971) sought to understand "the 'spontaneous' consent given by the great masses of the population to the general direction imposed on social life by the dominant fundamental

[p. 319 ↓] [p. 320 ↓] [p. 321 ↓] group." His analysis stressed the role of social influence and persuasion and distinguished between *active* (spontaneous) and *passive* forms of support for the social system (see also Hochschild, 1981: 260–283). Gramsci also recognized the potentially progressive or revolutionary role of ideology; that is, its ability to "organize" members of disadvantaged groups, to enable them to "acquire consciousness of their position," and to motivate what he (and other Marxists) regarded as a necessary "struggle." At the same time, Gramsci clearly perceived the ideological advantages possessed by those who side with the status quo:

Table 42.1 Theoretical precursors, part 1: approaches to dominant ideology and false consciousness

Author(s)	Source(s)	Key concepts	Illustrative quotations
Karl Marx and Friedrich Engels	<i>The German Ideology</i> (1846)	Ruling ideas, ideology, illusion, palliative function of religion	"The ideas of the ruling class are in every epoch the ruling ideas ... The class which has the means of material production at its disposal, has control at the same time over the means of mental production" "If in all ideology men and their circumstances appear upside-down as in a camera obscura, this phenomenon arises just as much from their historical life-process as the inversion of objects on the retina does from their physical life-process" "The people are interested in maintaining the present state of production" "Religion ... is the opium of the people" "To call [on the people] to give up their illusions about their condition is to call on them to give up a condition that requires illusions"

Author(s)	Source(s)	Key concepts	Illustrative quotations
Antonio Gramsci	<i>Selections from the Prison Notebooks</i> (1929–1935)	Dominant ideology, cultural hegemony, "spontaneous consent"	"[The] functions of social hegemony and political government [include]: The 'spontaneous' consent given by the great masses of the population to the general direction imposed on social life by the dominant fundamental group; this consent is 'historically' caused by the prestige (and consequent confidence) which the dominant group enjoys because of its position and function in the world of production" "To the extent that ideologies are historically necessary they have a validity which is 'psychological'; they 'organise' human masses, and create the terrain on which men move, acquire consciousness of their position, struggle, ..." "[A] popular conviction often has the same energy as a material force" "The existing social order is presented as a stable, harmoniously coordinated system, and the great mass of people hesitate and lose heart when they think of what a radical change might bring ... They can only imagine the present being torn to pieces, and fail to perceive the new order which is possible, and which would be better organized, more vital than the old one"
György Lukács	<i>History and Class Consciousness</i> (1918–1930)	Class consciousness, reification, false consciousness	"By relating consciousness to the whole of society it becomes possible to infer the thoughts and feelings which men would have in a particular situation if they were able to assess both it and the interests arising from it in their impact on immediate action and on the whole structure of society" "[The dialectical method does not permit us simply to proclaim the 'falseness' of this consciousness and to persist in an inflexible confrontation of true and false. On the contrary, it requires us to investigate this 'false consciousness' concretely as an aspect of the historical totality and as a stage in the historical process"
Jon Elster	"Belief, bias and ideology" (1982); <i>Sour Grapes</i> (1983)	Micro-foundations of ideological illusions, distortion, rationalization, "analytical Marxism"	"[There is a] tendency of the oppressed and exploited classes in a society to believe in the justice of the social order that oppresses them. This belief, perhaps, is mainly due to distortion, i.e., to such affective mechanisms as rationalization. But there is also an element of illusion, of bias stemming from purely cognitive sources" "The interest of the upper class is better served by the lower classes spontaneously inventing an ideology justifying their inferior status. This ideology, while stemming from the interest of the lower classes in the sense of leading to dissonance reduction, is contrary to their interest because of a tendency to overshoot, resulting in excessive rather than in proper meekness"

Author(s)	Source(s)	Key concepts	Illustrative quotations
Jennifer Hochschild	<i>What's Fair?</i> (1981)	Why doesn't "the dog bark"?	"[T]he American poor apparently do not support the downward distribution of wealth. The United States does not now have, and seldom ever has had, a political movement among the poor seeking greater economic equality. The fact that such a political movement could succeed constitutionally makes its absence even more startling. Since most of the population have less than an average amount of wealth – the median level of holdings is below the mean – more people would benefit than would lose from downward redistribution. And yet never has the poorer majority of the population, not to speak of the poorest minority, voted itself out of its economic disadvantage"
James R. Kluegel and Eliot R. Smith	<i>Beliefs about Inequality: Americans' Views of What Is and What Ought to Be</i> (1986)	Dominant ideology, stratification beliefs, psychological control, emotional benefits	"[C]ertain Marxist theories ... assume working-class people will come to recognize the contradictions between their self-interests and their system justifying beliefs" "[T]he belief in internal control, part of the dominant ideology, is adaptive for an individual's personal life. This belief leads to more positive and less negative emotional experience ... Psychological control – even if not always accompanied by real control of one's important life outcomes – seems to have positive consequences. These consequences in turn may be important in motivating people to maintain a belief in the dominant ideology as a whole"
Sandra L. Bem and Daryl J. Bem	"Case Study of a nonconscious ideology: training the woman to know her place" (1970)	Nonconscious ideology, illusion of equality, complementary stereotyping	"The ideological rationalization that men and women hold complementary but equal positions in society appears to be a fairly recent invention. In earlier times – and in more conservative company today – it was not felt necessary to provide the ideology with an egalitarian veneer" "In 1954 the United States Supreme Court declared that a fraud and hoax lay behind the slogan 'separate but equal.' It is unlikely that any court will ever do the same for the more subtle motto that successfully keeps the woman in her place: 'complementary but equal'"
Catharine A. MacKinnon	<i>Toward a Feminist Theory of the State</i> (1989)	Socialist-feminism, sexual objectification, consciousness raising	"Feminist inquiry ... begins with a broad unmasking through consciousness raising of the attitudes that legitimate and hide women's status, the daily practices and ideational envelope that contain woman's body" "Gender socialization is the ... process through which women internalize (make their own) a male image of their sexuality as their identity as women, and thus make it real in the world" "Male power is a myth that makes itself true. To raise consciousness is to confront male power in its duality: as at once total on one side and a delusion on the other. In consciousness raising, women learn they have learned that men are everything, women their negation, but the sexes are equal ... [N]o woman escapes the meaning of being a woman within a gendered social system"

The existing social order is presented as a stable, harmoniously coordinated system, and the great mass of people hesitate and lose heart when they think of what a radical change might bring ... They can only imagine the present being torn to pieces, and fail to perceive the new order which is possible, and which would be better organized, more vital than the old one.

(Gramsci, 1917, quoted in Fiori, 1970: 106–107)

As far as we know, the first expression of the term “system-justifying beliefs” occurs just once in a book by Kluegel and Smith (1986) entitled *Beliefs about Inequality*, which

combines sociological theory in the Gramscian tradition with an analysis of public opinion data. In passing, the authors refer to “certain Marxist theories that assume working-class people will come to recognize the *contradictions between their self-interests and their system-justifying beliefs*” (1986: 15; emphasis added). Kluegel and Smith were also apparently the first to demonstrate empirically that system-justifying beliefs are associated with emotional benefits, including a subjective sense of control, even among poor people. This notion, which is tacit in Marx's provocative claim that religious ideology is the “opium of the masses” (see Turner, 1991) and Lerner's (1980) formulation of the belief in a just world, inspired theoretical and empirical investigations of the *palliative* function of system-justifying beliefs and ideologies (Jost and Hunyady, 2002; Wakslak et al., 2007).

György Lukács, who was sometimes in the good graces of the Communist authorities in Hungary but often not, was in any case better positioned than Gramsci to explore the concept of *false consciousness*. Importantly, Lukács (1971: 59) recognized the necessity of distinguishing between subjective (or perceived) and objective (or actual) class interests:

By relating consciousness to the whole of society it becomes possible to infer the thoughts and feelings which men would have in a particular situation if they were able to assess both it and the interests arising from it in their impact on immediate action and on the whole structure of society.

Lukács also followed Marx in advocating an empirical, social scientific analysis of ideology and false consciousness, eschewing dogmatic, polemical, and purely speculative philosophical approaches (cf. Jost, 1995; Jost and Jost, 2007).

Elster (1982, 1983) explicitly advocated a scientific approach to Marxian hypotheses about the social functions of ideological beliefs (see also Cunningham, 1987; Runciman, 1969). This was part of a broader movement known as “analytical Marxism” (or, more colloquially, “non-Bullshit” Marxism), which sought to develop and test empirically falsifiable claims derived from Marxian theory on such topics as social class, exploitation, labor, value, and ideology. Here and elsewhere, Elster incorporated the work of social psychologists, including Festinger, Deutsch, Lerner, Abelson, Nisbett,

and Ross. In so doing, he advanced several hypotheses that reformulated Marxian ideas in unabashedly psychological terms, including these two statements, which directly influenced the formulation of system justification theory along motivational lines:

[There is a] tendency of the oppressed and exploited classes in a society to believe in the justice of the social order that oppresses them. This belief, perhaps, is mainly due to distortion, i.e., to such affective mechanisms as rationalization. But there is also an element of illusion, of bias stemming from purely cognitive sources. (Elster, 1982: 131)

The interest of the upper class is better served by the lower classes spontaneously inventing an ideology justifying their inferior status. This ideology, while stemming from the interest of the lower classes in the sense of leading to dissonance reduction, is contrary to their interest because of a tendency to overshoot, resulting in excessive rather than in proper meekness. (Elster, 1982: 142)

Thus, Elster argued that individuals' "ideological adaptation to their state of submission was endogenous, and not only did not require, but would have been incompatible with, deliberate ideological manipulation by the rulers" (1982: 124).

MacKinnon (1989) shrewdly critiqued Marxism from the perspective of feminism and feminism from the perspective of Marxism. The result was a sophisticated analysis of the social and political significance of "consciousness raising" activity, not only with respect to members of the working class but with respect to members of subjugated groups in general, including women. She argued that "no woman escapes the meaning of being a woman within a gendered social system" and that "male power becomes self-enforcing" to the extent that women internalize sexist norms and standards (1989: 99). MacKinnon's analysis comported with that of Bem and Bem (1970), who pointed out that sexist ideology can operate insidiously at a presumably nonconscious level of awareness (cf. Jost and Kay, 2005).

The notion that people are simultaneously embedded (and therefore psychologically invested) in multiple social systems and institutions, including capitalism, patriarchy, nuclear families, work organizations, and so on, is a crucial insight that inspired the

attempt to develop a general theory of system justification. The scientific goal, in other words, is not merely to explain the unique effects of any single system or institution (e.g., capitalism), but rather to identify general mechanisms or processes that play out in a wide variety of social systems (ranging from families to society as a whole) on the thoughts, feelings, and behaviors of individuals and groups (see Jost and Banaji, 1994; Wakslak et al., in press).

In an effort to synthesize various socialist–feminist approaches to the study of dominant ideology and false consciousness, Jost (1995) proposed the following (general) definition of the latter term: “The holding of false or inaccurate beliefs that are contrary to one's own social interest and which thereby contribute to the maintenance of the disadvantaged position of the self or the group” (1995: 400).¹ A literature review suggested the existence of at least six different types of false consciousness beliefs: (1) denial of injustice or exploitation, (2) fatalism about prospects for social change, (3) rationalization of social roles, (4) false attribution of blame, (5) identification with the oppressor, and (6) resistance to social change (Jost, 1995). Many authors, especially those influenced by postmodernist philosophy, have rejected the concept of false consciousness on the grounds that it is difficult (or even impossible) to distinguish between true and false statements in the social and political world. Without conceding this highly skeptical epistemological claim, we recognize that it may be a pragmatic advantage that the concept of system justification sidesteps the issue of whether beliefs and ideologies that sustain social systems are (wholly or partially) true or false. The focus, rather, is on their motivational bases and system-maintaining consequences.

Stereotyping, Prejudice, and the Internalization of Inferiority

From the start, system justification theory reflected an effort to unify the analysis of dominant ideology and false consciousness with social psychological research on stereotyping, prejudice, and the internalization of inferiority. Neither Kurt Lewin (1941) nor Gordon Allport (1954) could be considered Marxists, but both appreciated the deep extent to which hierarchical social systems impinge upon the thoughts, feelings,

and behaviors of the individual (see [Table 42.2](#)). Lewin, for instance, noted that “self-hatred is a phenomenon which [p. 323 ↓] [p. 324 ↓] [p. 325 ↓] has its parallel in many underprivileged groups,” including Jews and African Americans. As a field theorist, he clearly saw the problem in contextual, environmental (i.e., system-level) terms, writing that: “Jewish self-hatred will die out only when actual equality of status with the non-Jew is achieved.” Allport, too, identified the problem of internalization of inferiority and argued that, above all, stereotypes serve a “rationalizing and justifying function” (see also Eagly and Steffen, 1984; Hoffman and Hurst, 1990; Jost and Banaji, 1994). The German social theorist Norbert Elias distinguished between “established” and “outsider” groups in a wide range of social and cultural contexts and noted that “groups in an outsider position measure themselves with the yardstick of their oppressors” (Elias and Scotson, 1965: 26). Anti-colonial theorists such as Frantz Fanon (1961), Albert Memmi (1968), and Steven Biko (1978) concurred. Biko (1978), for instance, declared in the context of the Apartheid system in South Africa, “The most potent weapon in the hands of the oppressor is the mind of the oppressed.”

Table 42.2 Theoretical precursors, part 2: stereotyping, prejudice, and the internalization of inferiority

<i>Author(s)</i>	<i>Source(s)</i>	<i>Key concepts</i>	<i>Illustrative quotations</i>
Kurt Lewin	"Self-hatred among Jews" (1941); "Group decision and social change" (1947)	Group self-hatred, feelings of inferiority, resistance to change, field theory	<p>"Jewish self-hatred is a phenomenon which has its parallel in many underprivileged groups. One of the better known and most extreme cases of self-hatred can be found among American Negroes"</p> <p>"It is recognized in sociology that the members of the lower social strata tend to accept the fashions, values, and ideals of the higher strata"</p> <p>"Self-hatred seems to be a psychopathological phenomenon, and its prevention may seem mainly a task for the psychiatrist. However, modern psychology knows that many psychological phenomena are but an expression of a social situation in which the individual finds himself"</p> <p>"Jewish self-hatred will die out only when actual equality of status with the non-Jew is achieved"</p> <p>"The study of the conditions for change begins appropriately with an analysis of the conditions for 'no change,' that is, for the state of equilibrium"</p>
Gordon Allport	<i>The Nature of Prejudice</i> (1954)	Rationalizing function of stereotypes, internalization	<p>"[S]o heavy is the prevailing cultural pressure that members of minority groups sometimes look at themselves through the same lens as other groups"</p> <p>"[In some cases] the victim instead of pretending to agree with his 'betters' actually does agree with them and sees his own group through their eyes"</p> <p>"[T]he rationalizing and justifying function of a stereotype exceeds its function as a reflector of group attributes"</p> <p>"You and I are not normally aware of the extent to which our behavior is constrained and regulated by such features of the social system"</p>
Norbert Elias	<i>The Civilizing Process</i> (1939); <i>The Established and the Outsiders</i> (1965, with J. Scotson)	The established, outsiders, group charisma, stigmatization, power, superiority and inferiority, social figuration	<p>"Whether they are social cadres, such as feudal lords in relation to villeins, 'whites' in relation to 'blacks,' Gentiles in relation to Jews, Protestants in relation to Catholics and vice versa, men in relation to women (in former days), large and powerful nation-states in relation to others ... [or] an old-established working-class group in relation to members of a new working-class settlement in their neighborhood – in all these cases the more powerful groups look upon themselves as the 'better' people, as endowed with a kind of group charisma, with a specific virtue shared by all its members and lacked by the others. What is more, in all these cases the 'superior' people may make the less powerful people themselves feel that they lack virtue – that they are inferior in human terms"</p> <p>"[W]here the power differential is very great, groups in an outsider position measure themselves with the yardstick of their oppressors. In terms of their oppressors' norms they find themselves wanting; they experience themselves as being of lesser worth"</p>

Author(s)	Source(s)	Key concepts	Illustrative quotations
Frantz Fanon	<i>The Wretched of the Earth</i> (1961)	Colonization, psychology of the oppressed, national consciousness	<p>"The colonialist bourgeoisie is helped in its work of calming down the natives by the inevitable religion. All the saints who have turned the other cheek, who have forgiven trespasses against them, and who have been spat on and insulted without shrinking are studied and held up as examples"</p> <p>"The native is an oppressed person whose permanent dream is to become the persecutor. The symbols of social order – the police, the bugle calls in the barracks, military parades and the waving flags – are at one and the same time inhibitory and stimulating"</p> <p>"A belief in fatality removes all blame from the oppressor; the cause of misfortune and of poverty is attributed to God: He is Fate. In this way, the individual ... bows down before the settler and his lot, and by a kind of interior restabilization acquires a stony calm"</p>
Steven Biko	<i>I Write What I Like</i> (1978)	Colonialism, white domination, inferiority complex, "Black Consciousness"	<p>"[T]he most potent weapon in the hands of the oppressor is the mind of the oppressed"</p> <p>"[B]lacks are suffering from inferiority complex – a result of 300 years of deliberate oppression, denigration and derision"</p> <p>"All in all the black man has become a shell, a shadow of man, completely defeated, drowning in his own misery, a slave, an ox bearing the yoke of oppression with sheepish timidity"</p> <p>"[B]lack consciousness' seeks to show the black people the value of their own standards and outlook. It urges black people to judge themselves according to these standards and not to be fooled by white society who have white-washed themselves and made white standards the yardstick by which even black people judge each other"</p> <p>"We are all oppressed by the same system"</p>
Henri Tajfel	<i>Human Groups and Social Categories</i> (1981); "The social identity theory of intergroup relations, acceptance of the status quo, absence of 'cognitive alternatives'	Perceived stability and legitimacy of intergroup relations, acceptance of the status quo, absence of "cognitive alternatives"	<p>"Outgroup social stereotypes tend to be created and widely diffused in conditions which require: (i) a search for the understanding of complex and usually distressful, large-scale social events; (ii) justification of actions, committed or planned, against outgroups"</p> <p>"Where social-structural differences in the distribution of resources have been institutionalized, legitimized, and justified through a consensually accepted status system (or at least a status system that is sufficiently firm and pervasive to prevent the creation of cognitive alternatives to it), the result has been less and not more ethnocentrism in the different status groups"</p> <p>"[An] important requirement of research on social justice would consist of establishing in detail the links between social myths and the general acceptance of injustice"</p>

Tajfel (1981) elaborated a group-based version of Allport's argument that a stereotype's "function is to justify (rationalize) our conduct in relation to that category" (1954: 191). Specifically, he noted that stereotypes justify "actions, committed or planned, against outgroups." Tajfel stopped short of recognizing the *system-justifying* functions of stereotyping and prejudice (as distinct from their *group-justifying* functions), but he did see a correlation between perceptions of the social system and intergroup attitudes:

Where social-structural differences in the distribution of resources have been institutionalized, legitimized, and justified through a consensually accepted status system (or at least a status system that is sufficiently firm and pervasive to prevent the creation of cognitive alternatives to

it), the result has been less and not more ethnocentrism in the different status groups.

Researchers subsequently explored the phenomenon of outgroup favoritism among members of low status or disadvantaged groups (see Hinkle and Brown, 1990), which Jost and Banaji (1994) regarded as a likely manifestation of false consciousness (see also Jost, 1995). System justification theory was developed in order to better understand the social psychological processes whereby social systems are “institutionalized, legitimized, and justified;” that is, the ways in which people who occupy quite different statuses or positions in society nevertheless find reasons to embrace the status quo with an enthusiasm that may seem puzzling in retrospect or when viewed from historical distance.

The Tolerance of Social Injustice

Many social historians, including Chalmers Johnson (1966), Howard Zinn (1968), and Barrington Moore Jr. (1978), have observed that social stability and acceptance of injustice are far more common than protest and rebellion (see [Table 42.3](#)). The question is *why*. Whereas sociologists, political scientists, and many laypersons assume that unjust social orders are maintained by force or the threat of force, psychologists since at least the advent of cognitive dissonance theory (Elster, 1982; Festinger, 1957) have understood that people are capable of rationalizing even their own suffering (see also Henry and Saul, 2006; Jost et al., 2003c). Deutsch (1974, 1985) integrated Anna Freud's ideas about “identification with the aggressor” with Lewin's account of “group self-hatred” to explain why it is so difficult to “awaken the sense of injustice.” Social justice researchers have often developed psychological theories of why people tolerate injustice and deprivation; processes of denial, rationalization, and social comparison have all been implicated (e.g., Crosby et al., 1989; Major, 1994; Olson and Hafer, 2001). As Major (1994) noted, “One of the more intriguing phenomena of social justice is that people tend to legitimate the status quo, even when it is disadvantageous to the self.”

It was Lerner's (1980) account of the “belief in a just world,” however, that first postulated [p. 326 ↓] [p. 327 ↓] [p. 328 ↓] a sweeping motivational tendency (a

“fundamental delusion”) to believe that the social world is orderly, predictable, and above all *just*. Lerner pondered the oft-noted human propensity to imbue social regularities (what “is”) with an “ought” quality. He wrote: “People want to and have to believe they live in a just world so that they can go about their daily lives with a sense of trust, hope, and confidence in their future.” There are however, significant differences between just world and system justification perspectives (Jost and Hunyady, 2002; Jost et al., 2004b). Most notably, Lerner theorized that the “justice motive” would lead people first to attempt to help innocent victims and to rectify injustice; only when they are *prevented* from doing so did Lerner think that people would engage in denial, rationalization, and victim-blaming strategies to maintain the belief in a just world. System justification theory, by contrast, holds that people are motivated to *exaggerate* the fairness and desirability of social, economic, and political institutions and arrangements; such tendencies are assumed to be *antithetical* to the genuine desire to attain social justice in practice (Jost and Kay, 2010). According to system justification theory, people may blame victims and defend the status quo even when opportunities to rectify injustice are potentially available and – in a departure from cognitive dissonance theory – even when they bear no personal responsibility for aversive outcomes (see Jost et al., 2003c; Kay et al., 2002).

Table 42.3 *Theoretical precursors, part 3: the tolerance of social injustice*

Author(s)	Source(s)	Key concepts	Illustrative quotations
Chalmers Johnson	<i>Revolutionary Change</i> (1966)	Social stability and change, societal equilibrium and disequilibrium	"[P]eople in societies are not inherently mutinous. Society is a form of human interaction that transcends violence, of which one form is revolution. Revolutions are in this sense antisocial, testifying to the existence of extraordinary dissatisfactions ... They do not occur randomly, and they need not occur at all. Revolution can be rationally contemplated only in a society that is undergoing radical structural change and that is in need of still further change" "[T]he tragedy of the American race situation was that both Negroes and whites accepted a stable, envalued definition of Negro inferiority – and consequent role assignments – for most of the century after the Civil War... the main body of the Negro population did not support innovations developed by marginal men from their own group"
Howard Zinn	<i>Disobedience and Democracy</i> (1968)	"Congealed injustice," law and order, dissent	"Society's tendency is to maintain what has been. Rebellion is only an occasional reaction to suffering in human history; we have infinitely more instances of forbearance to exploitation, and submission to authority, than we have examples of revolt. Measure the number of peasant insurrections against the centuries of serfdom in Europe – the millennia of landlordism in the East; match the number of slave revolts in America with the record of those millions who went through their lifetimes of toil without outward protest. What we should be most concerned about is not some natural tendency toward violent uprising, but rather the inclination of people, faced with an overwhelming environment, to submit to it"
Bartington Moore Jr.	<i>Injustice: The Social Bases of Obedience and Revolt</i> (1978)	Suffering, oppression, explanation, justification, "psychological anesthesia"	"People are evidently inclined to grant legitimacy to anything that is or seems inevitable no matter how painful it may be. Otherwise the pain might be intolerable" "The human capacity to withstand suffering and abuse is impressive, tragically so" "[A]t the level of cultural norms and shared perceptions it will be necessary to overcome the illusion that the present state of affairs is just, permanent, and inevitable" "As any oppressed or suffering group seeks to come to terms with its fate, its members, and more especially its leaders and spokesmen, seek an explanation for that fate" "In a stratified society the principles of social inequality, generally systematized by priests, explain and justify the more prevalent and routine forms of suffering"

Author(s)	Source(s)	Key concepts	Illustrative quotations
Morton Deutsch	"Awakening the sense of injustice" (1974); <i>Distributive Justice</i> (1985)	Tolerance of injustice, identification with the aggressor, group self-hatred	"Although the need to maintain a positive self-regard is common, it is not universal. The victim of injustice, if he views himself favorably, may be outraged by his experience and attempt to undo it; in the process of doing so, he may have to challenge the victimizer. If the victimizer is more powerful than he and has the support of the legal and other institutions of the society, he will realize that it would be dangerous to act on his outrage or even to express it. Under such circumstances, in a process that Anna Freud (1937) labeled 'identification with the aggressor,' the victim may control his dangerous feelings of injustice and outrage by denying them and by internalizing the derogatory attitudes of the victimizer toward himself... Thus, he will become in Lewin's terms (1935) a 'self-hater' who attributes blame for his victimization upon himself or his group"
Melvin Lerner	<i>The Belief in a Just World: A Fundamental Delusion</i> (1980)	Belief in a just world, need for control, deservingness, myth, victim-blaming	"People want to and have to believe they live in a just world so that they can go about their daily lives with a sense of trust, hope, and confidence in their future" "As human beings, we judge events in moral terms. People, acts, outcomes are not only evaluated on some dimension of desirability; they are also viewed in terms of their 'appropriateness,' and we want it all to fit together in the appropriate way" "It is virtually a cliché in our culture to consider the poverty-stricken, or even the relatively deprived, as having their own compensating rewards. They are actually happy in their own way – carefree, happy-go-lucky, in touch with and able to enjoy the 'simple pleasures of life'... Some systems of religious belief see virtue in suffering, and assume restitution in later life"
Tom R. Tyler and Kathleen McGraw	"Ideology and the interpretation of personal experience: procedural justice and political quiescence" (1986)	Dominant ideology, cultural socialization, procedural fairness beliefs	"[T]he development of dysfunctional views about procedure is the result of cultural socialization. The disadvantaged accept the dominant American ideology of a 'contest' mobility system. Because they accept this procedure as fair, citizens conclude that the social allocation system is fair and do not examine the distributive consequences of this system; i.e., whether its outcomes actually correspond to distributive justice principles... By accepting societal values, the disadvantaged are led to focus upon aspects of their situation that are ineffective in inducing a sense of injustice and, hence, lead to political quiescence"

Author(s)	Source(s)	Key concepts	Illustrative quotations
Brenda Major	"From social inequality to personal entitlement: the role of social comparisons, legitimacy appraisals, and group membership" (1994)	Social comparison biases, depressed entitlement, legitimizing attributions	"Often, people who are objectively disadvantaged report themselves to be just as content and satisfied with their circumstances as are people who are objectively more privileged" "Social comparison biases tend to prevent awareness of disadvantage, and attribution biases tend to legitimize disadvantage" "People typically feel they deserve the same treatment or outcomes that they have received in the past or that others like themselves receive" "[E]ven when members of disadvantaged groups do become aware of their disadvantaged status, they often appraise it as legitimate. As a result, the disadvantaged often come to believe that they are personally entitled to less than do members of more advantaged groups" "One of the more intriguing phenomena of social justice is that people tend to legitimate the status quo, even when it is disadvantageous to the self" "[U]nequal social distributions have a powerful tendency to be legitimated"

The Institutional Legitimation of the Social Order

Whereas psychologists tend to emphasize the role of individual thoughts, feelings, and behaviors in the perpetuation of the status quo, sociologists (as a general rule) focus on the ways in which the social order accrues legitimacy by fostering social stability through cultural and ideological as well as coercive (i.e., social control) processes (see [Table 42.4](#)). For instance, Berger and Luckmann (1966) in *The Social Construction of Reality* noted that “institutions, once formed, have a tendency to persist.” To a considerable extent, this is because people [p. 329 ↓] [p. 330 ↓] accept “social reality” (the shared assumptions and understandings that are encountered in childhood and afterward) as “taken for granted.” Legitimacy, in other words, is the default assumption when people think and speak about the societal status quo. The needs of the individual (e.g., for “cohesion,” or epistemic coherence) gratify the needs of the system (for legitimacy) and vice versa.

Table 42.4 Theoretical precursors, part 4: the institutional legitimation of the social order

Author(s)	Source(s)	Key concepts	Illustrative quotations
Peter L. Berger and Thomas Luckmann	<i>The Social Construction of Reality</i> (1966)	Institutionalization, legitimization, need for cohesion, "taken-for-granted" nature of reality	"[T]he institutional world requires legitimization, that is ways by which it can be 'explained' and justified" "Institutionalization is not, however, an irreversible process, despite the fact that institutions, once formed, have a tendency to persist" "It is possible that this tendency to integrate meanings is based on a psychological need, which may in turn be physiologically grounded (that is, that there maybe a built-in 'need' for cohesion in the psycho-physiological constitution of man)" "The legitimization of the institutional order is also faced with the ongoing necessity of keeping chaos at bay"
Jürgen Habermas	<i>Legitimation Crisis</i> (1975); <i>The Theory of Communicative Action</i> (1987); <i>The New Conservatism</i> (1989)	Need for legitimization, justification, system crisis, ideology as "systematically distorted communication"	"In societies organized around a state, a need for legitimization arises that, for structural reasons, could not yet exist in tribal societies" "Social systems adapt inner nature to society with the help of normative structures in which needs are interpreted and actions licensed or made obligatory. The concept of motivation that appears here should not conceal the specific fact that social systems accomplish the integration of inner nature through the medium of norms that have need of justification" "[W]ithin the framework of a legitimate order of authority, the opposition of interests can be kept latent and integrated for a certain time. This is the achievement of legitimizing world-views or ideologies" "As soon ... as belief in the legitimacy of an existing order vanishes, the latent force embedded in the system of institutions is released"
Pierre Bourdieu	<i>Outline of a Theory of Practice</i> (1977)	Legitimating discourses, class habitus, "the established order"	"Once a system of mechanisms has been constituted capable of objectively ensuring the reproduction of the established order of its own motion ... the dominant class have only to let the system they dominate take its own course in order to exercise their domination" "[T]he task of legitimating the established order does not fall exclusively to the mechanisms traditionally regarded as belonging to the order of ideology, such as law ... The most successful ideological effects are those which have no need of words, and ask no more than complicitous silence" "[T]he ideological use many societies make of the lineage model and, more generally, of genealogical representations, in order to justify and legitimate the established order ... would doubtless have become apparent to anthropologists at an earlier date if the theoretical use they themselves make of this theoretical construct had not prevented them from inquiring into the functions of genealogies and genealogists"

Author(s)	Source(s)	Key concepts	Illustrative quotations
Vaclav Havel	"The power of the powerless" (1991)	Power, guilt, responsibility, social order	"Position in the power hierarchy determines the degree of responsibility and guilt, but it gives no one unlimited responsibility and guilt, nor does it completely absolve anyone. Thus the conflict between the aims of life and the aims of the system is not a conflict between two socially defined and separate communities; and only a very generalized view (and even that only approximative) permits us to divide society into the rulers and the ruled ... In the post-totalitarian system [the line of conflict] runs de facto through each person, for everyone in his own way is both a victim and a supporter of the system. What we understand by the system is not, therefore, a social order imposed by one group upon another, but rather something which permeates an entire society and is a factor in shaping it"
Mary R. Jackman	<i>The Velvet Glove</i> (1994)	Benevolent paternalism, ideology, coercion, social control	"[G]roups who enjoy the fruits of domination ... work to engage subordinates in a common view of the world that rationalizes the current order. The surest method of social control is to induce subordinates to regulate themselves. To that end, the unmediated weapon of choice is ideology" "Institutions can legitimize and stabilize inequality by removing compliance from the self-conscious realm ... [T]he advantage of authority over the explicit assertion of power is that the threat remains implicit, submerged beneath an elaborate ideological edifice"

The philosopher Jürgen Habermas (1975) also addressed the "need for legitimization" of social systems (e.g., the state) and the ways in which "social systems adapt inner [human] nature to society with the help of normative structures." In a semi-Marxian vein, Habermas noted that social systems experience "legitimation crises" and that "value consensus" is attained only when certain social conflicts are sufficiently repressed or kept out of sight. To the extent that *social integration* is achieved in this manner, value consensus is inherently "ideological" in the classic Marxian sense and requires justification, such as that offered by conservative ideology (Habermas, 1989; see also Jost et al., 2003b).

Bourdieu (1977) was similarly concerned with how relations of dominance and submission in society are sustained over time. In addition to formal mechanisms, such as the law, Bourdieu noted that "[t]he most successful ideological effects are those which ... ask no more than complicitous silence." As suggested also by Berger and Luckmann (1966), among others, it is possible to lend legitimacy to the status quo simply by going along with it; that is, *by appearing not to challenge it* (cf. Jost and Major, 2001). What these [p. 331 ↓] various sociological perspectives contribute to system justification theory is the notion that it is remarkably easy for social systems to enjoy legitimacy and stability by winning the apparent (if not always the actual) consent of the majority of the populace. To the extent that people consciously or unconsciously adapt

themselves to the systems that affect them, it seems that they cannot help but reinforce those systems.

Vaclav Havel, the first President of the Czech Republic, made a similar set of observations about Communist society in the late twentieth century:

In the post-totalitarian system [the line of conflict] runs de facto through each person, for everyone in his own way is both a victim and a supporter of the system. What we understand by the system is not, therefore, a social order imposed by one group upon another, but rather something which permeates an entire society and is a factor in shaping it.

Havel, 1978

Whereas other perspectives in social psychology (especially social identity and social dominance theories) tend to assume that the “social order is imposed by one group upon another,” system justification theory takes seriously Havel’s (1978) insight that individuals are active (as well as passive) participants in the justification (or legitimization) of the status quo, so that “everyone in his own way is both a victim and a supporter of the system.”

Although Mary Jackman (1994) rejects the concept of false consciousness and assumes that social actors (including the disadvantaged) are motivated by self-interest, she astutely accounts the myriad ways in which ideology is used as a “weapon” to insure “social control” by “rationalizing the current order.” In analyzing historical cases as diverse as the enslavement of Africans in the Old and New Worlds, and the role of women under patriarchy, Jackman shows how social stability results from a kind of *collaboration* between members of dominant and subordinate groups. She notes that flattering treatment (or “benevolent paternalism”) – as exemplified by men’s polite but chauvinistic tendency to place women “on a pedestal” (cf. Bem and Bem, 1970; Glick and Fiske, 2001) – helps to gain compliance. The notion that even flattering stereotypes of disadvantaged group members can be used to justify the status quo by ostensibly compensating for their state of disadvantage and creating an “illusion of equality” was

picked up and demonstrated experimentally in later research on system justification theory (Jost and Kay, 2005; Kay and Jost, 2003).

Conservatism, Authoritarianism, and Social Dominance

Writing in 1899, the sociologist Thorstein Veblen sought to understand the sociological and psychological bases of conservative ideology (see also Mannheim, 1925, 1936, and [Table 42.5](#)). Specifically, Veblen saw “an instinctive revulsion at any departure from the accepted way of doing and of looking at things – a revulsion common to all men and only to be overcome by stress of circumstances.” Half a century later, Adorno et al. (1950) identified conservatism with:

[A]n attachment, on the surface at least, to ‘things as they are,’ to the prevailing social organization and ways. Related to the idea that ‘what is, is right,’ is a tendency to idealize existing authority and to regard the ‘American way’ as working very well. Social problems tend either to be ignored or to be attributed to extraneous influences rather than to defects intrinsic in the existing social structure. One way of rationalizing chronic problems is to make them ‘natural’ ... Or, as a prominent ultra-conservative radio commentator observed recently: ‘There is nothing wrong with our American system. It is as good as it ever was, but we must do all we can in the New Year to get rid of the charlatans, fakers, and agitators who are responsible for so many problems.’ It is clear from the other speeches of this commentator that his ‘charlatans’ are for the most part leaders of the labor movement or of liberal political groupings – men who, in his eyes, threaten the existing order

(1950: 153–154)

Adorno and his colleagues associated extremely conservative (or what they termed “pseudoconservative”) outlooks with a propensity to [p. 332 ↓] [p. 333 ↓] follow authoritarian, even potentially fascist opinion leaders.

Table 42.5 Theoretical precursors, part 5: authoritarianism, social dominance, and political ideology

Author(s)	Source(s)	Key concepts	Illustrative quotations
Thorstein Veblen	<i>The Theory of the Leisure Class</i> (1899)	Habit, conservatism, aversion to change	"The opposition of the [wealthy] class to changes in the cultural scheme is instinctive, and does not rest primarily on an interested calculation of material advantages; it is an instinctive revulsion at any departure from the accepted way of doing and of looking at things – a revulsion common to all men and only to be overcome by stress of circumstances. All change in habits of life and of thought is irksome." "Any change in men's views as to what is good and right in human life makes its way but tardily at the best. Especially is this true of any change in the direction of what is called progress"
Theodor W. Adorno, Else Frankel-Brunswik, Daniel J. Levinson, and R. Nevitt Sanford	<i>The Authoritarian Personality</i> (1950)	Authoritarianism, fascist potential, ethnocentrism, pseudo-conservatism, resistance to social change	"Perhaps the definitive component of conservatism is an attachment, on the surface at least, to 'things as they are,' to the prevailing social organization and ways. Related to the idea that 'what is, is right,' is a tendency to idealize existing authority and to regard the 'American way' as working very well. Social problems tend either to be ignored or to be attributed to extraneous influences rather than to defects intrinsic in the existing social structure. One way of rationalizing chronic problems is to make them 'natural' ... To be 'liberal,' on the other hand, one must be able actively to criticize existing authority. The critics may take various forms, ranging from mild reforms (e.g., extension of government controls over business) to complete overthrow of the status quo"
Erich Fromm	<i>Beyond the Chains of Illusion</i> (1962)	Illusion, religious mystification, liberation	"The assumption underlying Marx's 'weapon of truth' is the same as with Freud: that man lives with illusions because these illusions make the misery of real life bearable" "'False consciousness,' that is to say, the distorted picture of reality, weakens man. Being in touch with reality, having an adequate picture of it, makes him stronger"
Robert E. Lane	"The fear of equality" (1959); <i>Political Ideology</i> (1962)	Rationalization, ideology, working class conservatism, "equality of happiness"	"The greater the emphasis in a society upon the availability of 'equal opportunity for all,' the greater the need for members of that society to develop an acceptable rationalization for their own social status" "The greater the strain on a person's self-esteem implied by a relatively low status in an open society, the greater the necessity to explain this status as 'natural' and 'proper' in the social order. Lower status people generally find it less punishing to think of themselves as correctly placed by a just society than to think of themselves as exploited, or victimized by an unjust society" "[I]t is as important to explain why revolutions and radical social movements do not happen as it is to explain why they do"

Author(s)	Source(s)	Key concepts	Illustrative quotations
Philip Mason	"Patterns of Dominance" (1971)	Dominance, psychological dependence, the "premise of inequality"	<p>"A social system based on inequality has to provide some degree of psychological satisfaction. Such systems could not otherwise have been so widespread nor have lasted so long"</p> <p>"It is to the rulers' interest that the subjects should think of the rulers as so different from themselves that they can never hope to supplant them, but also as 'their' rulers whom they must defend against outsiders"</p> <p>"The fear of reprisals is not the only force that keeps the slaves, the serfs, the peasants, the workers, subservient ... They must somehow be led to believe that the system is part of the order of nature and that things will always be like this"</p> <p>"That so many people for so much of history have accepted treatment manifestly unfair must always be puzzling to an observer from an individualist society"</p>
Jim Sidanius and Felicia Pratto	"Social dominance orientation: a personality variable predicting social and political attitudes" (1994); "Social Dominance" (1999)	Social dominance, hierarchy-enhancing, and hierarchy-attenuating legitimizing myths	<p>"Ideologies that promote or maintain group inequality are the tools that legitimize discrimination. To work smoothly, these ideologies must be widely accepted within a society, appearing as self-apparent truths: hence we call them hierarchy-legitimizing myths":</p> <p>"Despite significant variations in the degree of oppression from one society to another ... many societies share the basic social-psychological elements that contribute to inequality: socially shared myths that define 'superior group' and 'inferior group' and that attempt to justify this distinction and the policies that 'should' follow from it"</p> <p>"[W]ithin relatively stable group-based hierarchies, most of the activities of subordinates can be characterized as cooperative rather than subversive to the system of group-based domination"</p>

Erich Fromm, like Adorno and Habermas, was a member of the so-called Frankfurt School, which endeavored to merge Marx and Freud, in order to develop a deeper psychological (or psychoanalytic) basis for Marxist social theory. Fromm (1962) saw that both Marx and Freud, in their own ways, were battling against religious and other ideological illusions that kept people from fulfilling their potential. He also suggested in *Escape from Freedom* that most (if not all) human beings possess a fear of personal autonomy that makes them susceptible to authoritarian manipulation (Fromm, 1941). Lane (1959) provided a kindred psychodynamic analysis of working-class conservatives, many of whom were thought to harbor an unconscious "fear of equality." In a statement that foreshadows several key tenets of system justification theory, Lane (1959: 49) wrote:

The greater the strain on a person's self-esteem implied by a relatively low status in an open society, the greater the necessity to explain this status as 'natural' and 'proper' in the social order. Lower status people generally find it less punishing to think of themselves as correctly placed by a just society than to think of themselves as exploited, or victimized by an unjust society.

[p. 334 ↓]

Mason (1971) observed that “psychological dependence” often results from relations of domination and subordination. He noted that members of dominant groups frequently persuade subordinates to protect them from both internal and external threats to their hegemony. This is facilitated by fostering a sense of dependence and inevitability among subordinates (see [Table 42.5](#)).

Although Jost and Banaji did not learn about social dominance theory until the basic tenets of system justification theory had already been developed, the work of Sidanius and Pratto (1999) served as an inspiration and a foil. Both theoretical perspectives emphasize the ways in which ideologies and other belief systems serve to imbue the existing social order with legitimacy. They also concur in the judgment that “most of the activities of subordinates can be characterized as cooperative rather than subversive to the system of group-based domination.” However, system justification theory does not assume, as social dominance theory does, that natural selection pressures have created a strong, potentially insurmountable preference for unequal over equal social relations in human beings. Rather, system justification theory suggests that people are motivated and psychologically equipped to accept and justify a wide range of social systems.

Major Postulates of System Justification Theory and Illustrative Evidence

System justification theory cheerfully adheres to two basic laws of psychology as described by McGuire: “[F]irst, that basically everybody is the same; and second, that everybody is fundamentally different” (1980: 180). These are the first two postulates of the theory, as summarized in [Box 42.1](#). That is, system justification theory holds first and foremost that people in general are motivated (often unconsciously, i.e., without deliberate awareness or intention) to defend, justify, and bolster aspects of the status quo, including existing social, economic, and political institutions and arrangements (Postulate I). The general cognitive–motivational process of system justification, in other words, is expected to be largely the same for members of advantaged and

disadvantaged groups. Evidence does indicate that people from all walks of life and from many different cultural backgrounds appear to engage in system justification, at least to some extent (e.g., Henry and Hardin, 2006; Henry and Saul, 2006; Jost et al., 2005; Sibley et al., 2007; Ullrich and Cohrs, 2007; Yoshimura and Hardin, 2009). Thus, acceptance and maintenance of the status quo has been observed among the rich and poor, men and women, old and young, heterosexuals and homosexuals, as well as people of diverse national, ethnolinguistic, and racial backgrounds (Glick and Fiske, 2001; Henry and Saul, 2006; Jost and Kay, 2005; Jost et al., 2003c, 2004a; Kay and Jost, 2003; Kilianski and Rudman, 1998; Lau et al., 2008). Members of both advantaged and disadvantaged groups internalize rather than reject existing hierarchies, often favoring the advantaged over the disadvantaged on implicit and sometimes even explicit measures (e.g., Ashburn-Nardo and Johnson, 2008; Ashburn-Nardo et al., 2003; Jost, 2001; Jost et al., 2002, 2004a; Rudman et al., 2002; Uhlmann et al., 2002).

At the same time, the strength of system justification motivation and its expression are expected to vary according to situational (contextual) and dispositional (individual difference) factors (Postulate II). More specifically, system justification motivation is increased when the status quo is perceived to be (a) inevitable or inescapable, or (b) criticized, challenged, or threatened and when (c) the individual feels dependent on or controlled by the system or its representatives (Postulate III). For instance, several experiments reveal that threats to the legitimacy of the social system lead people to increase their use of stereotypes to justify social and economic inequalities and to defend the status quo more vigorously (Jost et al., 2005; Kay et al., 2005). Lau et al. (2008) extended this basic finding to the context of interpersonal attraction, demonstrating that system [p. 335 ↓] threat leads men to prefer female romantic partners who confirm sexist, system-justifying stereotypes over those who do not. In the U.S., evidence of heightened nationalism, patriotism, and increased support for governmental institutions and authorities (e.g., police, military, Congress, and the President) following the terrorist attacks of September 11, 2001, is consistent with the notion that system threat activates or enhances system justification motivation (Jost et al., 2010; Ullrich and Cohrs, 2007). Similarly, when people feel that restrictive emigration policies prevent them from leaving the country or system or when they otherwise feel extremely dependent upon the current system and its representatives,

they exhibit heightened system justification tendencies (e.g., Laurin et al., 2010; van der Toorn et al., 2011).

Box 42.1 Postulates of System Justification Theory

- I. People in general are motivated (often unconsciously, that is, without deliberate awareness or intention) to defend, justify, and bolster aspects of the status quo, including existing social, economic, and political institutions and arrangements.
- II. The strength of system justification motivation and its expression are expected to vary according to situational (contextual) and dispositional (individual difference) factors.
- III. System justification motivation is increased when the status quo is perceived to be (a) inevitable or inescapable, or (b) criticized, challenged, or threatened, and when (c) the individual feels dependent on or controlled by the system (or its representatives).
- IV. System justification satisfies basic epistemic motives to reduce uncertainty, existential motives to manage threat, and relational motives to coordinate social relationships. Thus, dispositional and situational variability in such needs will affect the strength of system justification motivation.
- V. There are several possible means by which the system can be justified, including direct endorsement of certain ideologies, the legitimization of institutions and authorities, denial or minimization of system problems or shortcomings, complementary stereotyping, rationalization, etc.
- VI. For members of advantaged groups (or those who are favored by the status quo), system justification is consonant with ego and group justification motives; it is therefore positively associated with self-esteem, ingroup favoritism, and long-term psychological wellbeing.
- VII. For members of disadvantaged groups (or those who are disfavored by the status quo), system justification conflicts with ego and group justification motives; it is therefore negatively associated with self-esteem, ingroup favoritism, and long-term psychological wellbeing.

- VIII. System justification serves a palliative function; that is, the endorsement of system-justifying beliefs and ideologies is associated in the short term with increased positive affect and decreased negative affect for members of advantaged and disadvantaged groups alike.
- IX. Although system justification motivation typically leads people to resist social change (and to perceive it as threatening to the status quo), people are more willing to embrace change when it is perceived as (a) inevitable or extremely likely to occur, and/or (b) congruent with the preservation of at least some aspects of the social system and/or its ideals.

System justification is thought to satisfy basic *epistemic* needs for consistency, certainty, and meaning; *existential* needs to manage threat and distress; and *relational* needs to coordinate social relationships and achieve shared reality with others (Postulate IV; see Jost et al., 2008a). It follows that dispositional and situational variability in such needs should affect the strength of system justification motivation (Jost and Hunyady, 2005). Consistent with this formulation, a meta-analysis by Jost et al. (2003b) revealed that uncertainty avoidance, intolerance of ambiguity, personal needs for order, structure, and closure, perceptions of a dangerous world, and death anxiety are all positively associated with an affinity for politically conservative, system-justifying (versus liberal, system-challenging) ideologies (see also Jost et al., 2007). Cognitive complexity, openness to new experiences, and the motivation to prolong cognitive closure on the other hand, are negatively associated with [p. 336 ↓] political conservatism and other forms of system justification (Carney et al., 2008).

Consistent with a goal systems approach to system justification (Jost et al., 2008b), there are several possible means by which the system can be justified, such as direct endorsement of certain ideologies, the legitimization of institutions and authorities, denial or minimization of system problems or shortcomings, rationalization, and so on (Postulate V). In the context of intergroup relations, system justification needs are frequently satisfied through processes of stereotyping, whereby members of both advantaged and disadvantaged groups accept and perpetuate the existing hierarchy by judging the advantaged to be more competent and industrious (and sometimes better overall) than the disadvantaged (Jost and Hunyady, 2002; Jost et al., 2005; Oldmeadow and Fiske, 2007; Sidanius and Pratto, 1999) and by endorsing

complementary stereotypes that create an “illusion of equality” in society (Bem and Bem, 1970; Jost and Kay, 2005; Kay and Jost, 2003).

There are a number of pre-existing *ideologies* or belief systems that people can embrace to bolster the societal status quo (Blasi and Jost, 2006). These include the Protestant work ethic, belief in a just world, meritocratic ideology, fair market ideology, economic system justification, power distance, benevolent sexism, social dominance orientation, right-wing authoritarianism, religious fundamentalism, and political conservatism (Glick and Fiske, 2001; Jost and Burgess, 2000; Jost and Thompson, 2000; Jost et al., 2003a, 2003b; Kaiser and Pratt-Hyatt, 2009; Major, 1994; Oldmeadow and Fiske, 2007; Sibley et al., 2007; Sidanius and Pratto, 1999). What these various belief systems have in common is that they explain social, economic, or political outcomes in a manner that generally maintains the subjective legitimacy of the status quo (Jost and Hunyady, 2005). Because of this feature, their endorsement should satisfy epistemic, existential, and relational needs to a greater extent than belief systems that are openly critical, contemptuous, or challenging of the status quo (e.g., Marxism, feminism, anarchy, and other revolutionary ideologies).² The fact that there are so many different types of system-justifying belief systems highlights the fact that such concerns permeate individuals' social relationships, family dynamics, and work lives, as well as their attitudes about society, religion, politics, economics, business, and the law. Research to date suggests that stereotyping, ideological endorsement, and other ways of justifying the system in various domains provide more or less interchangeable means of attaining the system justification goal in practice (Jost et al., 2010; Kay et al., 2005).

Some long-term social psychological consequences of system justification are theorized to be opposite for members of advantaged and disadvantaged groups (Jost and Thompson, 2000). For members of advantaged groups (or those who are favored by the status quo), the perception of being “on top” of society is consonant with the holding of positive attitudes concerning their own group and themselves (Jost et al., 2001, 2002). That is, system justification is consonant with ego and group justification motives; it is positively associated with self-esteem, ingroup favoritism, and psychological wellbeing (Postulate VI). Members of disadvantaged groups (or those who are disfavored by the status quo), however, are faced with a potential conflict between their need to justify

the system and competing motives to enhance their own self-esteem and group status. For them, system justification conflicts with ego and group justification motives; it is negatively associated with self-esteem, ingroup favoritism, and long-term psychological wellbeing (Postulate VII). Specifically, the more they justify the system (as well as their own group), the more disadvantaged group members exhibit ambivalence toward the ingroup-outgroup favoritism, and suffer in terms of subjective wellbeing as indexed by levels of self-esteem, neuroticism, depression, and generalized anxiety (Ashburn-Nardo et al., 2003, 2008; Jost and Burgess, 2000; Jost and Thompson, 2000; Jost et al., 2002; O'Brien and Major, 2005).

[p. 337 ↓]

At the same time, Jost and Hunyady (2002) proposed that system justification can serve a short-term palliative function for members of advantaged and disadvantaged groups alike. That is, the endorsement of system-justifying beliefs and ideologies is associated with increased positive affect, decreased negative affect, and satisfaction with the status quo (Postulate VIII). For instance, the tendency to embrace meritocratic ideology (e.g., believing that economic inequality is legitimate and necessary in capitalist society) is associated with increased life satisfaction and contentment among the poor as well as the rich (Jost et al., 2003c; Kluegel and Smith, 1986). The adoption of system-justifying beliefs and ideologies can reduce feelings of uncertainty, distress, guilt, frustration, helplessness, cognitive dissonance, and moral outrage brought on by social inequality and other potential system deficiencies (e.g., Kay et al., 2008; Wakslak et al., 2007).

Furthermore, the observed gap between liberals and conservatives in terms of self-reported happiness is at least partially explained by the latter's tendency to regard economic inequality as fair and just (Napier and Jost, 2008). It appears that system justification serves the palliative function of increasing positive affect and reducing negative affect for many of its adherents, although once again racial and ethnic minorities may benefit less (Rankin et al., 2009). Nevertheless, the occurrence of system justification is probably better explained by the fact that it helps to satisfy underlying epistemic, existential, and relational needs and not simply because it makes people feel "good" (cf. Elster, 1982).

Most system justification research conducted to date paints a fairly bleak picture with regard to prospects for social change. System justification, like rationalization in general, does seem to help people cope with unwelcome realities (Jost and Hunyady, 2002; Kay et al., 2002), but it also hampers the remediation of injustice and other system-level problems. Studies carried out by Wakslak et al. (2007) demonstrated that because system justification reduces moral outrage, it undermines the implementation of intentions and actions designed to help the disadvantaged (e.g., support for soup kitchens, job training programs, crisis hotlines, minority outreach, and after-school tutoring programs). Thus, system justification motivation typically leads people to resist social change and to perceive it as threatening to the status quo (Postulate IX).

At the same time, however, research suggests that people are more willing to embrace social change when it is perceived as (a) inevitable or extremely likely to occur, and/or (b) congruent with the preservation of at least some aspects of the social system and/or its ideals. For example, people engage in anticipatory rationalization of a new or emerging regime as soon as its implementation is seen as inevitable, or at least highly probable (Kay et al., 2002). One would expect this process to be facilitated by a rapid transition that replaces the previous regime entirely, thereby avoiding the problem of divided loyalties between current and former systems. Feygina et al. (2010) demonstrated that although system justification tendencies are generally associated with greater denial of global climate change and less commitment to proenvironmental action, it is possible to eliminate the negative effect of system justification on environmentalism by framing proposed changes as “system-sanctioned” (i.e., as patriotic and consistent with protecting the status quo). With the right kind of leadership and message framing, then, it may be possible to harness system justification motivation in a constructive manner so that it enables people to improve upon the status quo rather than reflexively defending against the possibility of change.

Practical Implications of System Justification Theory

Robert Lynd (1939) suggested that “the role of the social sciences is to be troublesome, to disconnect the habitual arrangements by which we manage to live along, and to demonstrate the possibility of change in more adequate [p. 338 ↓] directions.” Conceived of in this way, a primary task of the social scientist is to overcome his or her system-justifying tendencies in order to see society as it really is, so that genuine problems can be recognized and, ultimately, resolved. As Lynd anticipated, “the role of such a constructive troublemaker is scarcely inviting,” presumably because it requires confronting and questioning the system-justifying assumptions of others. Indeed, system justification theorists (and, to a much more serious extent, many of the historical predecessors cited in [Tables 42.1–42.5](#)) have received more than usual levels of criticism; indeed, if the theory is sound, one would expect defensive, even aggressive reactions to the suggestion that societal arrangements are not as fair or legitimate as most people believe (see also Blasi and Jost, 2006). As Memmi observed, “[P]eople are always accused of exaggeration when they describe injustices to those who do not want to hear about them” (1968: 19). This brings us to an important practical insight of system justification theory, namely that people are wont to ignore, deny, minimize, rationalize, or dismiss criticisms or putative shortcomings of the status quo, even if these – like the problems posed by global climate change – are grounded in scientific evidence.

There are many other practical implications of system justification theory, some of which we have touched on throughout the chapter. These include consequences for intergroup relations involving status or power differences; the pernicious effects of relatively subtle forms of sexism and stereotyping; and the persistence of implicit as well as explicit prejudice. There are also implications of system justification theory for more overtly political behavior, including voting preferences, evaluations of system leaders and representatives, and the psychological advantages conferred by incumbency status and conservative ideology. System justification theory has been used to elucidate the effects of public policies such as restricted immigration (Laurin et al., 2010) and

to explain an ever-widening range of consequential outcomes, including religious commitment (Kay et al., 2008), romantic preferences (Eastwick et al., 2009; Lau et al., 2008), academic performance (Chatard et al., 2008), and willingness to help the disadvantaged (Wakslak et al., 2007). In the spirit of divergent theorizing, one hopes that by applying system justification theory to more and more domains, an increasingly refined set of scientifically and practically significant conclusions will emerge.

Conclusion

In this chapter we have traced the historical development of system justification theory, which offers a unique social psychological framework for understanding why so many individuals assume, even in the face of contrary evidence, that the societal status quo is, good, fair, legitimate, desirable, and right. We have identified nine major postulates of the theory and discussed how the theory builds on and extends the work of philosophers, historians, sociologists, psychologists, and political scientists. By positing that system justification is a motivated, goal-directed process for most of the people at least some of the time, we have offered a distinctive psychological perspective that helps to explain, among other things, the perceived legitimacy and stability of social systems and hierarchical arrangements, as well as the relative scarcity of protest and rebellion. Research will continue to identify moderating and mediating variables that help to explain when and why people will reflexively defend, bolster, and maintain the current social system and when they will not. Future work will speak even more directly to the ways in which resistance to change and political acquiescence can be transformed into an open, restless, critical, constructive search for forms of social organization that are better, truer, freer, more sustainable, and just. While these are grand scientific and practical pretensions for a single theory, the ideas have been simmering for at least a century and a half – indeed, since the advent of social science itself.

[p. 339 ↓]

Acknowledgments

The authors would like to thank Rick Andrews, Raynee Gutting, Erin Hennes, P.J. Henry, György Hunyady, Lawrence Jost, Aaron Kay, Arie Kruglanski, Alison Ledgerwood, Anesu Mandisodza, Avital Mentovich, Artur Nilsson, Andrew Shipley, Paul van Lange, and Jesse Wynhausen for superb comments on earlier drafts. The writing of this chapter was supported by the National Science Foundation (Award # BCS–0617558).

Notes

1 In the Marxian tradition, it is possible for members of advantaged groups to hold false beliefs that are *congruent* with self-interest and that still count as instances of false consciousness (e.g., Lukács, 1971). Because it is unclear in these cases, from a social psychological perspective, whether such beliefs are *motivated* by ego, group, or system justification needs, the clearest, least ambiguous cases of false consciousness (and system justification motivation) come from members of disadvantaged groups, who could not simply be motivated by self-interest or group-interest to defend and bolster the societal status quo (Jost, 1995; Jost and Banaji, 1994).

2 To the extent that revolutionaries engage in justification of an anticipated (i.e., utopian) status quo and they are convinced that the transition is inevitable, it is conceivable that revolutionary ideology could satisfy epistemic, existential, and relational motives to some degree (see also Blasi and Jost, 2006). At the same time, it seems self-evident that revolutionary activity itself requires one to embrace uncertainty, danger, and the risk of losing more than a few friends. It follows, that revolutionaries are primarily motivated by other, opposing needs or concerns (e.g., for justice, innovation, group justification, hatred, excitement, etc.).

John T.Jost Jojanekkevan der Toorn

References

- Abercrombie, N., ed. , Hill, S., ed. and Turner, B. (eds) (1990) Dominant Ideologies . London: Unwin Hyman.
- Adorno, T.W., Frenkel-Brunswik, E., Levinson, D.J. and Sanford, R.N. (1950) The Authoritarian Personality . New York: Harper.
- Allport, G. (1954) The Nature of Prejudice . Cambridge, MA: Addison-Wesley.
- Ashburn-Nardo, L. and Johnson, N. Implicit outgroup favoritism and intergroup judgment: The moderating role of stereotype context . Social Justice Research , 21, 490–508. (2008) <http://dx.doi.org/10.1007/s11211-008-0078-8>
- Ashburn-Nardo, L., Knowles, M.L. and Monteith, M.J. Black Americans' implicit racial associations and their implications for intergroup judgment . Social Cognition , 21, 61–87. (2003) <http://dx.doi.org/10.1521/soco.21.1.61.21192>
- Bargh, J.A., Gollwitzer, P.M., Lee Chai, A., Barndollar, K. and Trötschel, R. The automated will: Nonconscious activation and pursuit of behavioral goals . Journal of Personality and Social Psychology , 81, 1014–1027. (2001) <http://dx.doi.org/10.1037/0022-3514.81.6.1014>
- Bem, S.L. and Bem, D.J. (1970) Case study of a nonconscious ideology: Training the woman to know her place . In D.J. Bem (ed.), Beliefs, Attitudes, and Human Affairs , pp. 89–99. Belmont, CA: Brooks/Cole.
- Berger, P.L. and Luckmann, T. (1966) The Social Construction of Reality: A Treatise in the Sociology of Knowledge . Garden City, NY: Anchor.
- Biko, S. (1978) I Write What I Like . London: Heinemann.
- Blasi, G. and Jost, J.T. System justification theory and research: Implications for law, legal advocacy, and social justice . California Law Review , 94, 1119–1168. (2006) <http://dx.doi.org/10.2307/20439060>

Borges, J.L. (1964) Kafka and his precursors . In J.L. Borges (ed.), Other Inquisitions 1937–1952 . Austin: University of Texas Press.

Bourdieu, P. (1977) Outline of a Theory of Practice (Edition, 1986) New York: Cambridge University Press.

Carney, D.R., Jost, J.T., Gosling, S.D. and Potter, J. The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind . *Political Psychology* , 29, 807–840. (2008) <http://dx.doi.org/10.1111/j.1467-9221.2008.00668.x>

Chatard, A., Selimbegovic, L. and Konan, P. Leftists' and rightists' IQ as a function of stereotype salience . *Journal of Research in Personality* , 42, 602–1606. (2008) <http://dx.doi.org/10.1016/j.jrp.2008.05.004>

Crosby, F.J., Pufall, A., Snyder, R.C., O'Connell, M. and Whalen, P. (1989) The denial of personal disadvantage among you, me, and all the other ostriches . In M. Crawford, ed. and M. Gentry (eds), *Gender and Thought: Psychological Perspectives* , pp. 79–99. New York: Springer. http://dx.doi.org/10.1007/978-1-4612-3588-0_5

Cunningham, F. (1987) False consciousness . In *Democratic Theory and Socialism* , pp. 236–267. Cambridge: Cambridge University Press.

Deutsch, M. (1974) Awakening the sense of injustice . In M. Lerner, ed. , and M. Ross (eds), *The Quest for Justice* , pp. 19–42. New York: Holt.

Deutsch, M. (1985) Distributive Justice: A Social-psychological Perspective . New Haven: Yale.

Eagly, A.H. and Steffen, V.J. Gender stereotypes stem from the distribution of women and men into social roles . *Journal of Personality and Social Psychology* , 46, 735–754. (1984) <http://dx.doi.org/10.1037/0022-3514.46.4.735>

Eastwick, P.W., Richeson, J.A., Son, D. and Finkel, E.J. Is love colorblind? Political orientation and interracial romantic desire . *Personality and Social Psychology Bulletin* , 35, 1258–1268. (2009) <http://dx.doi.org/10.1177/0146167209338524>

Elias, N. and Scotson, J.L. (1965) *The Established and the Outsiders: A Sociological Enquiry into Community Problems* (2nd Edition, 1994). London: Sage. <http://dx.doi.org/10.4135/9781446222126>

Elster, J. (1982) Belief, bias, and ideology . In M. Hollis, ed. and S. Lukes (eds), *Rationality and Relativism* , pp. 123–148. Cambridge, MA: MIT Press.

Elster, J. (1983) Sour Grapes: Studies in the Subversion of Rationality . Cambridge: Cambridge University Press.

Fanon, F. (1961) *The Wretched of the Earth* (Edition, 2001). Harmondsworth: Penguin.

Festinger, L. (1957). *A Theory of Cognitive Dissonance* . Stanford, CA: Stanford University Press.

Feyginal., Jost, J.T. and Goldsmith, R. System justification, the denial of global warming, and the prospect of ‘system-sanctioned’ change . *Personality and Social Psychology Bulletin* , 36, 326–338. (2010). <http://dx.doi.org/10.1177/0146167209351435>

Fiori, G. (1970). *Antonio Gramsci: Life of a Revolutionary* . London: Verso.

Fromm, E. (1941) *Escape From freedom* (Edition, 1994). New York: Owl.

Fromm, E. (1962) *Beyond the Chains of Illusion: My Encounter with Marx and Freud* . New York: Simon and Schuster.

Fromm, E. (ed.) (1965) *Socialist Humanism: An International Symposium* . Garden City, NY: Doubleday.

Glick, P. and Fiske, S.T. An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality . *American Psychologist* , 56, 109–118. (2001) <http://dx.doi.org/10.1037/0003-066X.56.2.109>

Gramsci, A. (1971) *Selections from the Prison Notebooks* . New York: International Publishers.

- Gurr, T. (1970) Why Men Rebel . Princeton, NJ: Princeton University Press.
- Habermas, J. (1975) Legitimation Crisis . Boston, MA: Beacon.
- Habermas, J. (1989) The New Conservatism: Cultural Criticism and the Historians' Debate . Cambridge, MA: MIT Press.
- Hardin, C.D. and Higgins, E.T. (1996) Shared reality: How social verification makes the subjective reality . In R.M. Sorrentino, ed. and E.T. Higgins (eds), *Handbook of Motivation and Cognition* , 3, 28–84. New York: Guilford Press.
- Havel, V. (1978) The power of the powerless (abridged) . In V. Havel (ed.), *Open Letters* (Edition, 1991), pp. 127–145. London: Faber.
- Henry, P.J. and Hardin, C.D. The contact hypothesis revisited: Status bias in the reduction of implicit prejudice in the United States and Lebanon . *Psychological Science* , 17, 862–868. (2006) <http://dx.doi.org/10.1111/j.1467-9280.2006.01795.x>
- HenryP.J. and SaulA. The development of system justification in the developing world . *Social Justice Research* , 19, 365–378. (2006) <http://dx.doi.org/10.1007/s11211-006-0012-x>
- Hinkle, S. and Brown, R. (1990) Intergroup comparisons and social identity: Some links and lacunae . In D. Abrams, ed. and M.A. Hogg (eds), *Social Identity Theory : Constructive and Critical Advances* , pp. 48–70. London: Harvester.
- Hirschman, A.O. (1970) Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States . Cambridge, MA: Harvard University Press.
- Hochschild, J.L. (1981) What's Fair? American Beliefs About Distributive Justice . Cambridge, MA: Harvard University Press.
- Hoffman, C. and Hurst, N. Gender stereotypes: Perception or rationalization? *Journal of Personality and Social Psychology* , 58, 197–208. (1990) <http://dx.doi.org/10.1037/0022-3514.58.2.197>

Jackman, M.R. (1994) *The Velvet Glove: Paternalism and Conflict in Gender, Class, and Race Relations*. Berkeley: University of California Press.

Johnson, C. (1966) *Revolutionary Change* (2nd Edition, 1983). Stanford, CA: Stanford.

Jost, J.T. Negative illusions: Conceptual clarification and psychological evidence concerning false consciousness . *Political Psychology* , 16, 397–424. (1995) <http://dx.doi.org/10.2307/3791837>

Jost, J.T. (2001) Outgroup favoritism and the theory of system justification: An experimental paradigm for investigating the effects of socio-economic success on stereotype content . In G. Moskowitz (ed.), *Cognitive Social Psychology: The Princeton Symposium on the Legacy and Future of Social Cognition* , pp. 89–102. Mahwah, NJ: Erlbaum.

Jost, J.T. and Banaji, M.R. The role of stereotyping in system-justification and the production of false consciousness . *British Journal of Social Psychology* , 33, 1–27. (1994) <http://dx.doi.org/10.1111/j.2044-8309.1994.tb01008.x>

Jost, J.T., Banaji, M.R. and Nosek, B.A. A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo . *Political Psychology* , 25, 881–919. (2004a) <http://dx.doi.org/10.1111/j.1467-9221.2004.00402.x>

Jost, J.T., Blount, S., Pfeffer, J. and Hunyady, G. Fair market ideology: Its cognitive-motivational underpinnings . *Research in Organizational Behavior* , 25, 53–91. (2003a) <http://dx.doi.org/10.1016/S0191-3085%2803%2925002-4>

Jost, J.T. and Burgess, D. Attitudinal ambivalence and the conflict between group and system justification motives in low status groups . *Personality and Social Psychology Bulletin* , 26, 293–305. (2000) <http://dx.doi.org/10.1177/0146167200265003>

Jost, J.T., Burgess, D. and Mosso, C. (2001) Conflicts of legitimization among self, group, and system: The integrative potential of system justification theory . In J.T. Jost, ed. and B. Major (eds), *The Psychology of Legitimacy: Emerging Perspectives on Ideology*,

Justice, and Intergroup Relations , pp. 363–388. New York: Cambridge University Press.

Jost, J.T., Fitzsimons, G.M. and Kay, A.C. (2004b) The ideological animal: A system justification view . In J. Greenberg, ed. , S.L. Koole, ed. , and T. Pyszczynski (eds), Handbook of Experimental Existential Psychology , pp. 263–282. New York: Guilford Press.

Jost, J.T., Glaser, J., Kruglanski, A.W. and Sulloway, F. Political conservatism as motivated social cognition . Psychological Bulletin , 129, 339–375. (2003b) <http://dx.doi.org/10.1037/0033-2909.129.3.339>

Jost, J.T. and Hunyady, O. The psychology of system justification and the palliative function of ideology . European Review of Social Psychology , 13, 111–153. (2002) <http://dx.doi.org/10.1080/10463280240000046>

Jost, J.T. and Hunyady, O. Antecedents and consequences of system-justifying ideologies . Current Directions in Psychological Science , 14, 260–265. (2005) <http://dx.doi.org/10.1111/j.0963-7214.2005.00377.x>

Jost, J.T. and Kay, A.C. Exposure to benevolent sexism and complementary gender stereotypes: Consequences for specific and diffuse forms of system justification . Journal of Personality and Social Psychology , 88, 498–509. (2005) <http://dx.doi.org/10.1037/0022-3514.88.3.498>

Jost, J.T. and Kay, A.C. (2010) Social justice: History, theory, and research . In S.T. Fiske, ed. , D. Gilbert, ed. and G. Lindzey (eds), Handbook of Social Psychology, Vol. 2 (5th Edition), pp. 1122–1165. New York: Wiley.

Jost, J.T., Kivetz, Y., Rubini, M., Guermandi, G. and Mosso, C. System-justifying functions of complementary regional and ethnic stereotypes: Cross-national evidence . Social Justice Research , 18, 305–333. (2005) <http://dx.doi.org/10.1007/s11211-005-6827-z>

Jost, J.T., Ledgerwood, A. and Hardin, C.D. Shared reality, system justification, and the relational basis of ideological beliefs . *Social and Personality Psychology Compass* , 2, 171–186. (2008a) <http://dx.doi.org/10.1111/j.1751-9004.2007.00056.x>

Jost, J.T., Liviatan, I., Van der Toorn, J., Ledgerwood, A., Mandisodza, A. and Nosek, B. (2010) System justification: How do we know it's motivated? In D.R. Bobocel, ed. , A.C. Kay, ed. , M.P. Zanna, ed. , and J.M. Olson (eds), *The Psychology of Justice and Legitimacy: The Ontario Symposium* , 11, 173–203. Hillsdale, NJ: Erlbaum.

Jost, J.T., ed. and Major, B. (eds) (2001) *The Psychology of Legitimacy: Emerging Perspectives on Ideology, Justice, and Intergroup relations* . New York: Cambridge University Press.

Jost, J.T., Napier, J.L., Thorisdottir, H., Gosling, S., Palfai, T.P. and Ostafin, B. Are needs to manage uncertainty and threat associated with political conservatism or ideological extremity? *Personality and Social Psychology Bulletin* , 33, 989–1007. (2007) <http://dx.doi.org/10.1177/0146167207301028>

Jost, J.T., Pelham, B.W. and Carvallo, M.R. Nonconscious forms of system justification: Implicit and behavioral preferences for higher status groups . *Journal of Experimental Social Psychology* , 38, 586–602. (2002) <http://dx.doi.org/10.1016/S0022-1031%2802%2900505-X>

Jost, J.T., Pelham, B.W., Sheldon, O. and Sullivan, B.N. Social inequality and the reduction of ideological dissonance on behalf of the system: Evidence of enhanced system justification among the disadvantaged . *European Journal of Social Psychology* , 33, 13–36. (2003c) <http://dx.doi.org/10.1002/ejsp.127>

Jost, J.T., Pietrzak, J., Liviatan, I., Mandisodza, A.N. and Napier, J.L. (2008b) System justification as conscious and nonconscious goal pursuit . In J. Shah, ed. and W. Gardner (eds), *Handbook of Motivation Science* , pp. 591–605. New York: Guilford Press.

Jost, J.T. and Thompson, E.P. Group-based dominance and opposition to equality as independent predictors of self-esteem, ethnocentrism, and social policy attitudes

among African Americans and European Americans . Journal of Experimental Social Psychology , 36, 209–232. (2000) <http://dx.doi.org/10.1006/jesp.1999.1403>

Jost, L.J. and Jost, J.T. Why Marx left philosophy for social science . Theory and Psychology , 17, 297–322. (2007) <http://dx.doi.org/10.1177/0959354307075047>

Kaiser, C.R., Dyrenforth, P.S. and Hagiwara, N. Why are attributions to discrimination interpersonally costly? A test of system- and group-justifying motivations . Personality and Social Psychology Bulletin , 32, 1423–1536. (2006) <http://dx.doi.org/10.1177/0146167206291475>

Kaiser, C.R. and Pratt-Hyatt, J.S. Distributing prejudice unequally: Do whites direct their prejudice toward strongly identified minorities? Journal of Personality and Social Psychology , 96, 432–445. (2009) <http://dx.doi.org/10.1037/a0012877>

Kay, A.C., Gaucher, D., Napier, J.L., Callan, M.J. and Laurin, K. God and the government: Testing a compensatory control mechanism for the support of external systems . Journal of Personality and Social Psychology , 95, 18–35. (2008) <http://dx.doi.org/10.1037/0022-3514.95.1.18>

Kay, A.C., Gaucher, D., Peach, J.M., Friesen, J., Laurin, K., Zanna, M. and Spencer, S.J. Inequality, discrimination, and the power of the status quo: Direct evidence for a motivation to view what is as what should be . Journal of Personality and Social Psychology , 97, 421–434. (2009) <http://dx.doi.org/10.1037/a0015997>

Kay, A., Jimenez, M.C. and Jost, J.T. Sour grapes, sweet lemons, and the anticipatory rationalization of the status quo . Personality and Social Psychology Bulletin , 28, 1300–1312. (2002) <http://dx.doi.org/10.1177/01461672022812014>

Kay, A.C. and Jost, J.T. Complementary justice: Effects of ‘poor but happy’ and ‘poor but honest’ stereotype exemplars on system justification and implicit activation of the justice motive . Journal of Personality and Social Psychology , 85, 823–837. (2003) <http://dx.doi.org/10.1037/0022-3514.85.5.823>

Kay, A.C., Jost, J.T. and Young, S. Victim-derogation and victim-enhancement as alternate routes to system justification . Psychological Science , 16, 204–246. (2005) <http://dx.doi.org/10.1111/j.0956-7976.2005.00810.x>

Kilianski, S.E. and Rudman, L.A. Wanting it both ways: Do women approve of benevolent sexism? Sex Roles , 39, 333–352. (1998) <http://dx.doi.org/10.1023/A:1018814924402>

Klandermans, B. (1997) The Social Psychology of Protest . Oxford: Blackwell.

Kluegel, J.R. and Smith, E.R. (1986) Beliefs About Inequality: Americans' View of What Is and What Ought to Be . Hawthorne, NJ: Gruyter.

Lane, R.E. The fear of equality . American Political Science Review , 53, 35–51. (1959) <http://dx.doi.org/10.2307/1951729>

Lau, G.P., Kay, A.C. and Spencer, S.J. Loving those who justify inequality: The effects of system threat on attraction to women who embody benevolent sexist ideals . Psychological Science , 19, 20–21. (2008) <http://dx.doi.org/10.1111/j.1467-9280.2008.02040.x>

Laurin, K., Shepherd, S. and Kay, A.C. Restricted emigration, system inescapability, and defense of the status quo: System-justifying consequences of restricted exit opportunities . Psychological Science , 21, 1075–1082. (2010) <http://dx.doi.org/10.1177/0956797610375448>

Ledgerwood, A., Jost, J.T., Mandisodza, A.N. and Pohl, M. (in press) Working for the system: Motivated defense of meritocratic beliefs . Social Cognition .

Lerner, M.J. (1980) The Belief in a Just World: A Fundamental Delusion . New York: Plenum Press.

Lewin, K. (1941) Self-hatred among Jews . In K. Lewin (ed.), Resolving Social Conflicts (Edition, 1948), pp. 186–200. New York: Harper.

Lukács, G. (1971) Selections from History and Class Consciousness . In R.S. Gottlieb (ed.), An Anthology of Western Marxism: From Lukács and Gramsci to Socialist-Feminism (Edition, 1989). New York: Oxford University Press.

Lynd, R. (1939) Knowledge for What? Princeton, NJ: Princeton University Press.

MacKinnon, C.A. (1989) Toward a Feminist Theory of the State . Cambridge, MA: Harvard University Press.

Major, B. From social inequality to personal entitlement: The role of social comparisons, legitimacy appraisals, and group membership . Advances in Experimental Social Psychology , 26, 293–355. (1994) <http://dx.doi.org/10.1016/S0065-2601%2808%2960156-2>

Mannheim, K. (1925) Conservatism. A Contribution to the Sociology of Knowledge (Edition, 1986). London: Routledge.

Mannheim, K. (1936) Ideology and Utopia . London: Routledge.

Marcuse, H. (1964) One-dimensional Man: Studies in the Ideology of Advanced Industrial Society . Boston: Beacon.

Marx, K. and Engels, F. (1846) The German Ideology (Edition, 1970; C.J. Arthur, ed.). New York: International Publishers.

Mason, P. (1971) Patterns of Dominance . London: Oxford University Press.

McGuire, W.J. (1980) Interview with R.I. Evans (ed.), The Making of Social Psychology: Discussions with Creative Contributors , pp. 171–186. New York: Gardner.

McGuire, W.J. and McGuire, C.V. (1991) The content, structure, and operation of thought systems . In R.S. Wyer Jr., ed. and T.K. Srull (eds), Advances in Social Cognition , IV, 1–78. Hillsdale, NJ: Erlbaum.

Memmi, A. (1968) Dominated Man; Notes Towards a Portrait . New York: Orion.

Moore, B., Jr. (1978) Injustice: The Social Bases of Obedience and Revolt . White Plains, NY: Sharpe.

Moscovici, S. Notes towards a description of social representations . European Journal of Social Psychology , 18, 211–250. (1988) <http://dx.doi.org/10.1002/ejsp.2420180303>

Napier, J.L. and Jost, J.T. Why are conservatives happier than liberals? Psychological Science , 19, 565–572. (2008) <http://dx.doi.org/10.1111/j.1467-9280.2008.02124.x>

O'Brien, L.T. and Major, B. System-justifying beliefs and psychological well-being: The roles of group status and identity . Personality and Social Psychology Bulletin , 31, 1718–1729. (2005) <http://dx.doi.org/10.1177/0146167205278261>

Oldmeadow, J. and Fiske, S.T. System-justifying ideologies moderate status = competence stereotypes: Roles for belief in a just world and social dominance orientation . European Journal of Social Psychology , 37, 1135–1148. (2007) <http://dx.doi.org/10.1002/ejsp.428>

Olson, J.M. and Hafer, C.L. (2001) Tolerance of personal deprivation . In J.T. Jost, ed. and B. Major (eds), The Psychology of Legitimacy: Emerging Perspectives on Ideology, Justice, and Intergroup Relations , pp. 157–175. New York: Cambridge University Press.

Rankin, L.E., Jost, J.T. and Wakslak, C.J. System justification and the meaning of life: Are the existential benefits of ideology distributed unequally across racial groups? Social Justice Research , 22, 312–333. (2009) <http://dx.doi.org/10.1007/s11211-009-0100-9>

Reicher, S. The context of social identity: Domination, resistance, and change . Political Psychology , 20, 921–945. (2004) <http://dx.doi.org/10.1111/j.1467-9221.2004.00403.x>

Rudman, L.A., Feinberg, J. and Fairchild, K. Minority members' implicit attitudes: Automatic ingroup bias as a function of group status . Social Cognition , 20, 294–320. (2002) <http://dx.doi.org/10.1521/soco.20.4.294.19908>

Runciman, W. (1969) False consciousness . In W. Runciman (ed.), Sociology in Its Place , pp. 212–223. Cambridge: Cambridge University Press.

Sherif, M. (1936) The Psychology of Social Norms . Oxford: Harper.

Sibley, C.G., Overall, N.C. and Duckitt, J. When women become more hostilely sexist toward their gender: The system-justifying effect of benevolent sexism . *Sex Roles* , 57, 743–754. (2007) <http://dx.doi.org/10.1007/s11199-007-9306-1>

Sidanius, J. and Pratto, F. (1999) Social Dominance: An Intergroup Theory of Social Hierarchy and Oppression . New York: Cambridge University Press.

Tajfel, H. (1981) Human Groups and Social Categories . Cambridge, MA: Cambridge Press.

Turner, D. (1991) Religion: Illusions and liberation . In T. Carver (ed.), The Cambridge Companion to Marx , pp. 320–337. Cambridge: Cambridge University Press.

Turner, J.C. Tyranny, freedom and social structure: Escaping our theoretical prisons . *British Journal of Social Psychology* , 45, 41–46. (2006) <http://dx.doi.org/10.1348/014466605X79840>

Tyler, T.R. and Jost, J.T. (2007) Psychology and the law: Reconciling normative and descriptive accounts of social justice and system legitimacy . In A.W. Kruglanski, ed. and E.T. Higgins (eds), *Social Psychology: Handbook of Basic Principles* (2nd Edition), pp. 807–825. New York: Guilford Press.

Tyler, T.R. and McGraw, K.M. Ideology and the interpretation of personal experience: Procedural justice and political quiescence . *Journal of Social Issues* , 42, 115–128. (1986) <http://dx.doi.org/10.1111/j.1540-4560.1986.tb00228.x>

Uhlmann, E., Dasgupta, N., Elgueta, A., Greenwald, A.G. and Swanson, J. Subgroup prejudice based on skin color among Hispanics in the United States and Latin America . *Social Cognition* , 20, 198–226. (2002) <http://dx.doi.org/10.1521/soco.20.3.198.21104>

Ullrich, J. and Cohrs, J.C. Terrorism salience increases system justification: Experimental evidence . Social Justice Research , 20, 117–139. (2007) <http://dx.doi.org/10.1007/s11211-007-0035-y>

van der Toorn, J., Tyler, T.R. and Jost, J.T. More than fair: Outcome dependence, system justification, and the perceived legitimacy of authority figures . Journal of Experimental Social Psychology , 47, 127–138. (2011) <http://dx.doi.org/10.1016/j.jesp.2010.09.003>

Veblen, T. (1899) The Theory of the Leisure Class: An Economic Study of Institutions . New York: Heubsch.

Wakslak, C.J., Jost, J.T. and Bauer, P. (in press) Spreading rationalization: Increased support for large-scale and small-scale social systems following system threat . Social Cognition .

Wakslak, C.J., Jost, J.T., Tyler, T.R. and Chen, E.S. Moral outrage mediates the dampening effect of system justification on support for redistributive social policies . Psychological Science , 18, 267–274. (2007) <http://dx.doi.org/10.1111/j.1467-9280.2007.01887.x>

Yoshimura, K. and Hardin, C.D. Cognitive salience of subjugation and the ideological justification of U.S. geopolitical dominance in Japan . Social Justice Research , 22, 298–311. (2009) <http://dx.doi.org/10.1007/s11211-009-0102-7>

Zinn, H. (1968) Disobedience and Democracy: Nine Fallacies on Law and Order (Edition, 2002). Cambridge, MA: South End.

<http://dx.doi.org/10.4135/9781446249222.n42>

Handbook of Theories of Social Psychology

Justice Theory

Contributors: Tom R. Tyler

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Justice Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n43>

Print pages: 344-362

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n43>

[p. 344 ↓]

Chapter 43: Justice Theory

Abstract

Psychological research has repeatedly shown that people are motivated by more than their concern with maximizing gains and minimizing loses. They also want to do what is just, appropriate, and fair. In particular, people's thoughts feelings and actions are shaped by their sense of what is just. Two forms of justice are considered: distributive justice, which involves fair allocations, and procedural justice, the study of fair processes for making decisions. Within each area of justice one concern of psychologists is with research examining whether and when this form of justice influences people's thoughts, feelings, and behaviors. In both areas the study of justice suggests strongly that justice matters. However, while research consistently shows that justice matters, there are several psychological theories concerning why it matters. In particular, some researches emphasize the role of justice in facilitating social exchange, while others focus on the importance of justice for personal and group identities. I argue here that research generally supports an identity-based view of justice.

Introduction

We live in an historical era in which naïve social theories most often describe human nature as being dominated by people's efforts to pursue personal self-interest, a self-interest broadly defined in terms of material rewards and costs. In contrast to this widely held image, many areas of social psychology show that people have a larger set of motivations that are more strongly linked to values and that include concerns for the wellbeing of other people and for groups, communities, organizations, and societies (Tyler, *in press*; Van Lange et al., 2007). This includes studies of empathy, altruism, justice, moral values, and identity. The image of the person that we accept is

important because it speaks to the approach we take to a wide variety of social issues. It influences how we motivate people to cooperate, how we enforce rules, and how we address widely varied issues such as social welfare, healthcare, climate change, economic growth, and many others.

Within this set of literatures on social motivations, one of the most well developed and clearly specified set of psychological models concerns the psychology of justice. The first question asked by justice researchers is whether people care about justice. And a core contribution of the psychology of justice has [p. 345 ↓] been the demonstration that people's judgments, feelings, and behaviors in social settings are influenced by their views about what is just and fair, views that are distinct from evaluations of personal or group self-interest (Tyler and Smith, 1998). This literature indicates that people are not simply motivated by their assessments of what is beneficial or harmful to them, those they care about, or the groups, organizations, and societies to which they belong. People also make judgments about what is just or unjust, using criterion that reflect justice-based evaluations, and those judgments shape their reactions in social contexts (Tyler, 2000; Tyler and Smith, 1998; Tyler et al., 1997).

The occurrence, strength, and nature of this justice motivation influences how society should address social issues. For example: "What type of appeals are likely to be effective when we seek the redistribution of resources to those less well off?;" "How can we motivate people to care about the well being of future generations so that they conserve energy and work to combat global warming?;" "How do we allocate scarce collective resources when providing aid to different people and companies when trying to stimulate the economy?;" and "How do we manage divisive social issues such as abortion or gay marriage?" In each case, the most effective and viable approach to the issue depends upon our understanding of the psychology of justice.

Arenas of Justice Research

Justice theory argues that people are motivated by concerns about justice. However, within that general framework researchers have studied a variety of types of justice. This chapter will focus on two types of justice: distributive and procedural justice. It will not discuss retributive justice; that is, reactions to rule breaking (see Darley and

Pittman, 2003; Vidmar and Miller, 1980). In each of these areas of psychological research the psychology of justice explores judgments about the principles used to decide what is fair or unfair within social settings. Issues of justice have been important within psychology ever since the World War II era, a period during which there was an explosion of psychology theory and research into the study of social settings and group processes. During the same historical period, psychology also moved beyond psychological models that paid little attention to people's subjective evaluations of the world and the field became more concerned with how people interpreted their social experiences. Concerns about justice emerged once the importance of subjective assessments of social situations became clearly recognized within relative deprivation research (Tyler and Smith, 1998).

Distributive Justice

Theories of distributive justice tie comparisons to issues of justice (Walster et al., 1978). They do so by arguing that people compare their outcomes to standards of what is a fair or deserved outcome. In other words, people have a sense of what they are entitled to receive and evaluate their outcomes against this standard. This involves distributive justice – the fairness of the allocation of desirable outcomes across people. People express the greatest satisfaction when they receive a fair distribution, in comparison to receiving more or less in absolute terms but thinking that they are being given “too much” or “too little;” that they will incur material losses to pressure others to redistribute resources via principles of justice; and that they will leave situations they view as characterized by the unfair distribution of resources to move to situations where resource distribution is fairer, but in which they receive fewer rewards.

The central premise of distributive justice theories is that people react to what they receive in relation to what they deserve. As noted, there are two potentially unhappy groups: those who receive too little and those who receive too much. As might be expected, [p. 346 ↓] those who receive less than they feel they deserve are found to be angry and to engage in a variety of behaviors, ranging from working less to rioting. Justice researchers have studied many instances in which people have received less than they deserve, and have shown that this leads to strong negative emotional reactions and to efforts to seek restitution. Among disadvantaged groups, complex

psychological dynamics are unleashed because the disadvantaged often lack the power to compel justice and must therefore find ways to manage their feelings of unfairness.

Interestingly, and less intuitively, those who get too much are also found to be unhappy and to engage in efforts to either restore distributive justice by mechanisms such as working harder or giving resources away, or if those solutions are not practical, by leaving the situation. This latter distributive justice finding is especially important because it suggests that the desire to act fairly can influence the advantaged to take actions on behalf of less well off others.

Of course, while the distributive justice literature argues that people react to deviations from standards of fairness, that argument can be tested only if the standards being used to determine fairness can be determined. Morton Deutsch (1975) has presented three core principles of distributive justice: equity, equality, and need. Equality involves giving everyone similar outcomes, while equity and need differentiate among people either in terms of their productivity or their needs. Deutsch suggests that the use of each principle promotes different social goals: equity leads to productivity, equality to social harmony, and need to social welfare.

Procedural Justice

Procedural justice is the study of people's subjective evaluations of the justice of procedures – whether they are fair or unfair, ethical or unethical, and otherwise accord with people's standards of fair processes for social interaction and decision making. Procedural justice should be distinguished from subjective assessments of distributive justice. In most nontrivial situations some type of process is needed for gathering relevant evidence, deciding upon and implementing decision rules, and managing the interpersonal processes of gaining acceptance for allocations or of resolving conflicts. The area of procedural justice focuses on understanding the fairness of such processes.

Studies in this area show that peoples' choices among allocation procedures are influenced by their evaluations of their relative procedural fairness, as well as by the favorability and fairness of their outcomes; that peoples' satisfaction with and

willingness to accept allocations and dispute resolution decisions depends upon the fairness of the procedures used to make them; and that peoples' rule following behavior and cooperation with others are shaped by the procedural fairness of groups, organizations, and societies.

Subjective procedural justice judgments have been the focus of a great deal of research attention by psychologists because they have been found to be a key influence on a wide variety of important group attitudes and behaviors (Lind and Tyler, 1988; Tyler, 2000). Procedural justice has been especially important in studies of decision acceptance and rule following. One reason that people might comply with rules and authorities is that they receive desirable rewards for cooperating and/or fear sanctioning from the group for not cooperating. An alternative reason that people might comply is that they are motivated by their sense of justice to accept what they feel is fair, even if it is not what they want.

A key question is whether justice is effective in resolving conflicts and disagreements when people cannot have everything that they want. To the degree that people defer because allocation decisions are seen as fair, justice is an important factor in creating and maintaining social harmony. Research suggests that social justice can act as a mechanism for resolving social conflicts, and that [p. 347 ↓] procedural justice is especially central in such situations.

John Thibaut and Laurens Walker (1975) conducted the first experiments designed to show the impact of procedural justice. Their studies demonstrated that people's assessments of the fairness of third-party decision-making procedures predicted their satisfaction with outcomes. This finding has been widely confirmed in many subsequent laboratory and field studies of procedural justice, studies which show that when third-party decisions are fairly made people are more willing to voluntarily accept them. What is striking is that such procedural justice effects are widely found in studies of real disputes, in real settings, involving actual disputants and are found to have an especially important role in shaping adherence to agreements over time.

Summary

The different aspects of the psychological study of justice outlined are united by the finding that people are very sensitive to issues of justice in their dealings with other people in social settings. In fact, such justice-based judgments are found to be key drivers of a wide variety of reactions, including attitudes, emotions, and behaviors. John Rawls (1971) famously argued that “justice is the first virtue of social institutions” and the findings of psychological research on justice strongly support the parallel suggestion that people view justice as a pivotal evaluation shaping their relationships with one another. Hence, while people might react to their experiences in social settings in terms of personal self-interest, they do not. Instead, they react to their sense of what is just.

Further, when groups, organizations, or societies are seeking to organize themselves, they become centrally preoccupied with issues of justice. The ability of authorities and institutions to be viewed as legitimate and, consequently, to be able to call upon group members for voluntary cooperation is linked to whether they are viewed as just. The centrality of justice to the organization and functioning of society supports the suggestion that justice must be a key issue when people seek to resolve social problems or manage social conflicts. It is this connection between justice and the dynamics of groups that first led to my own interest in injustice. Anyone who studies political and social processes is soon struck by the centrality of “justice” to people's arguments, irrespective of their particular policy positions. Just as children rapidly learn to argue “that's not fair” societies frame their efforts to make allocations and resolve conflicts around struggles over the meaning of justice.

What Type of Justice Matters?

While research on distributive justice has provided evidence that people are most satisfied with fair outcomes in allocation, within group settings distributive justice has been found to be a problematic solution to problems of allocation. Studies indicate that distributive justice judgments are often the product of motivated social cognitions, with people's judgments about what they deserve shaped by the self-interested tendency to

exaggerate their contributions to collective products (Messick and Sentis, 1985; Ross and Sicoly, 1979; Thompson and Loewenstein, 1992). As a result, people often cannot be given what they feel they deserve, since their judgments of entitlement do not reflect the views of others about their actual contributions.

For this or other reasons, studies of justice in allocations and dispute resolution indicate that people focus less upon issues of distributive justice than they do upon two other issues: the procedures used to make allocation decisions and their interpersonal treatment within those procedures. These two issues have been collectively referred to as procedural justice. These findings emerge both from studies that look at the weight placed upon these different issues (Tyler and Caine, 1981) and from studies that look [p. 348 ↓] at what people talk about when asked to describe situations in which they feel that injustice has occurred (Messick et al., 1985; Mikula et al., 1990). In both of these types of studies the minimal role played by outcome distributions – real or perceived – in experiences of injustice is striking. Hence, in allocations people's focus is found to rest heavily on procedural and interpersonal issues.

Huo (2002) uses a different approach to addressing this issue, but reaches a similar conclusion. She creates a framework in which participants are asked about what should be given to a disliked group. Three issues are considered: monetary resources, procedural protections, and/or treatment with fairness and respect. Her results indicate that when people consider denying members of the disliked group and various things like denial of interpersonal treatment with dignity and respect is considered the most serious denial, while the denial of monetary resources is the least serious. Denial of procedural protections is intermediate. These findings suggest that people view procedural issues as more important than outcomes, with the quality of interpersonal treatment being especially central to the connection between people.

The conclusion of these studies comparing people's focus on different forms of justice is that people are most strongly affected by issues of procedural justice. And within procedural justice both the fairness of decision making and the quality of interpersonal treatment are found to have an influence upon people's reactions. Because it is central to people's concerns in their dealings with others this discussion will focus primarily on issues of procedural justice.

My own interest in procedural justice develops out of a concern with the dynamics of authority in groups, organizations, and societies. Groups respond to problems by creating rules, authorities, and institutions. The success of these organized social entities then becomes a key societal concern. That effectiveness is linked to the ability of those authorities and institutions to gain voluntary acceptance of their policies, rules, and decisions, a point made long ago by Lewin (Gold, 1999). And, being viewed as legitimate is central to such acceptance (Tyler, 2006a). Studies consistently show that procedural justice shapes the legitimacy of the authorities and institutions with which people deal, and through such attitudes, their willingness to defer to those authorities and institutions. Studies of the legitimacy of authority suggest that people decide how much to defer to authorities and to their decisions primarily by assessing the fairness of their decision-making procedures (Tyler, 2006).

Consequently, using fair decision-making procedures is the key to developing, maintaining, and enhancing the legitimacy of rules and authorities and gaining voluntary deference to social rules. Beyond issues of rule following, studies of procedural justice indicate that it plays an equally important role in motivating commitment to organizations. As a consequence, procedural justice is important in encouraging productivity and extra-role behavior in work organizations (Tyler and Blader, 2000, 2003). Procedural justice is a key antecedent of a wide variety of desirable cooperative behaviors in groups, organizations and societies (see Tyler, *in press*). Hence, procedural justice is central to any concern with effective group dynamics.

Defining Procedural Justice

From the beginning models of justice have drawn from broader social psychological models of the relationship between people and groups. The primary model that dominates early social psychology is social exchange (Thibaut and Kelley, 1959). This model argues that people's thoughts, feelings, and behaviors when dealing with others are guided by the desire to obtain material rewards and avoid material costs. This image is central to both theories of distributive justice (e.g., equity theory) and to Thibaut and Walker's control model of procedural justice (Thibaut and [p. 349 ↓] Walker, 1975). However, as research on justice has developed, it has increasingly been guided by a second model: social identity theory (Tajfel and Turner, 1979). This model

suggests that people use groups to create and support their identities. More recent models based upon social exchange have also become broader in scope and now recognize social preferences; for example, issues of egalitarianism and altruism (see Kelley and Thibaut, 1978; Kelley et al., 2003; Rusbult and Van Lange, 1996; Van Lange et al., 2007).

My own theoretical development reflects this gradual transition from social exchange perspectives to an emphasis on social identity. That transition reflects the findings of justice research, which are consistently at odds with the predictions of social exchange theories and are more consistent with a social identity framework. My research initial drew upon the social identity framework reflected in Lind and Tyler (1988), and has since involved to include an emphasis on intragroup status dynamics (i.e., respect as well as pride) and a more elaborate understanding of how identity dynamics operate that expands the idea of identity (DeCremer and Tyler, 2005). This shift in the focus of my work comes in response to the findings of studies on justice in real world settings. Tyler and Lind (1992), for example, is based upon field studies conducted among litigants. Litigant concerns were consistently found to be about issues that had very little to do with the gain or loss of resources. They were instead about being treated disrespectfully, about distrust in the motives of authorities and about feeling that those making decisions did not listen to and consider their concerns.

As noted, early work on procedural justice was guided by the influential control model of Thibaut and Walker (1975). Thibaut and Walker centered their procedural justice studies on procedures as mechanisms for settling disputes about the allocation of outcomes. Thibaut and Walker linked their discussions of procedures primarily to issues of decision-making, and in particular to issues of decision making about allocation decisions. Since their procedural models were rooted in an era where distributive justice dominated, their focus was natural. They also linked people's desire for fair procedures to their desire to achieve equitable outcomes. They proposed that people value procedural justice (i.e., voice or process control) because it facilitates decision-makers' ability to make equitable judgments. In other words, procedures are valued insofar as they affect the outcomes that are associated with them.

Thibaut and Walker (1975) focus upon a narrow set of antecedents of procedural justice, considering only questions of evidence presentation (process control) and

outcome control. However, other procedural justice researchers, in particular Leventhal (1980), identify a broader range of antecedents. In a wide-ranging discussion of procedural justice Leventhal identified six core procedural elements: consistency, bias-suppression, accuracy, correctability, representativeness, and ethicality. Of these, representativeness reflects the issues of voice central to Thibaut and Walker's model, while the other elements present a broader set of issues for evaluating procedures.

The exclusive focus on decision making in allocation contexts is no longer true of procedural justice research. Researchers have increasingly moved their attention away from an exclusive focus on the decision-making function of procedures to include more attention to the interpersonal aspects of procedures. Those interpersonal aspects of procedures arise because procedures are settings within which people are involved in a social interaction with one another. This is true irrespective of whether the procedure involves a decision maker.

In social interactions there is considerable variation in the manner in which people treat one another. They can act politely, rudely, respectfully, with hostility, and so on. These aspects of the interpersonal experience of a procedure – which occur in the context of an interaction whose overt purpose is to make a decision to allocate resources or resolve a conflict – may also influence those who are involved. These interpersonal aspects of [p. 350 ↓] procedures have been found by recent studies to be so powerful in their impact that some researchers have argued that they might potentially be treated as a separate type of “interactional” justice (Bies and Moag, 1986; Tyler and Bies, 1990). Irrespective of whether the quality of the treatment that people experience via procedures is actually considered a distinct form of justice (see Blader and Tyler, 2003), justice researchers have again followed their findings about what impacts the people they study. This has led them to increasingly turn their research toward exploring interpersonal or interactional aspects of procedures – for example, the quality of one's treatment by others.

Tyler and Lind (1992) drew upon Leventhal to develop the relational model of authority, which explores the role of this broader set of procedural factors in shaping reactions to authorities. They demonstrate that other factors – neutrality, trustworthiness, and status recognition – influence procedural justice judgments and shape reactions to authorities (Tyler, 1988, 1989; 1994; Tyler et al., 1996). Tyler (1994) further demonstrates that

procedural justice judgments are distinct from concerns about outcomes and outcome favorability.

Recent discussions of procedural justice recognize four elements of procedures as the primary factors that contribute to judgments about their fairness: opportunities for participation, a neutral forum, trustworthy authorities, and treatment with dignity and respect. Blader and Tyler (2003) refer to the first two elements as involving the quality of decision making, while the latter two elements are concerned with the quality of interpersonal treatment.

The voice effect indicates that people feel more fairly treated if they are allowed to participate in the resolution of their problems or conflicts. People are primarily interested in presenting their perspective and sharing in the discussion of conflicts that affect them, not in controlling decisions about how to handle such conflicts. Instead, people often look to authorities for resolutions. They expect authorities to make final decisions about how to act based upon the arguments those who are affected by those decisions have presented.

People are also influenced by judgments about neutrality – the honesty, impartiality, and objectivity of the authorities with whom they deal. They believe that authorities should not allow their personal values and biases to enter into their decisions, which should be made based upon rules and facts. Basically, people seek a “level playing field” in which no one is unfairly disadvantaged. If they believe that the authorities are following impartial rules and making factual, objective decisions, they think procedures are fairer.

Another factor shaping people's views about the fairness of a procedure is their assessment of the motives of the third-party authority responsible for resolving the case. People recognize that third parties typically have considerable discretion to implement procedures in varying ways, and they are concerned about the motivation underlying the decisions made by the authority with which they are dealing. Important assessments include whether that person is benevolent and caring, is concerned about their situation and their concerns and needs, considers their arguments, tries to do what is right for them, and tries to be fair.

Studies suggest that people also value having respect shown for their rights and for their status within society. They want their dignity as people and their rights as members of the society to be recognized and acknowledged. Surprisingly, such assessments of respect are largely unrelated to the outcomes they receive. Thus, the importance which people place upon this affirmation of their status is especially relevant to conflict resolution. Unlike the outcomes that determine distributive justice, dignity and respect is something that authorities can give to everyone with whom they deal.

The Interplay of Justice and Self-Interest

The studies of justice outlined make clear that justice can motivate people to behave [p. 351 ↓] in ways that are not in accord with their sense of their own personal and group interests. For example, the advantaged may give resources to the disadvantaged. On the other hand, such justice motivations are never absolute. Typically people compromise between the motivation to act justly and the desire to act in their self-interest.

One of the best illustrations of such compromises is found in the literature on ultimatum games (Handgraaf et al., 2004). In the ultimatum game, the proposers make offers about how to divide some set of resources. The responder can either accept or reject this offer. Studies suggest that proposers make, and responders accept, offers somewhere between an equal division and a division favoring the proposer. For example, if ten dollars is to be divided the successful offers fall between zero and five dollars. In other words, both parties compromise between self-interest and fairness, with the proposer giving more than they would be rationally expected to, and the responder accepting less than an equal division. Further, studies show that responders will decline small gains rather than accept “unfairly” low divisions illustrating that people are willing to incur losses to defend principles of fairness.

Of course, there are other ways to deal with conflicts about justice. Early work on distributive justice pointed to the possibility of motivated social cognition, that is, that people might try to restore justice psychologically (Walster et al., 1978). That distributive justice research first developed the distinction between psychological and behavioral responses to wrongdoing. When someone receives too much or provides too little

to others, a conflict is created between their behavior and the principles of justice. There are two types of response. One is for outcomes to be reallocated so as to be fair. The victim frequently advocates this response, while the harm doer has mixed feelings – they believe in justice but are also benefiting from the unjust situation. Hence, harmdoers are motivated to psychologically justify the situation, coming to believe that they deserve the outcomes they have. For example, studies of distributive justice show that people who are “overpaid” find ways to justify their payment by increasing their sense of the difficulty of the task, and hence reframing the situation as one in which their pay is reasonable.

The motivation to justify advantage brings harm doers and victims into conflict because the victim wants redistribution while the harm doer seeks to justify their gains. An important function of social authorities is to lend support to victims, or at least avoid social conflict, by supporting the application of objective standards of fairness, which resolves conflicts, and by discouraging psychological justification, which leads to long-term hostility. More generally, there are a variety of social mechanisms through which the advantaged justify their advantage, with the intention of keeping their advantages without the negative emotions that they experience from feeling that they are violating principles of justice (Chen and Tyler, 2001; Wakslak et al., 2007).

Most recently Blader (2007) has demonstrated that such motivational judgments occur when the justice of procedures is ambiguous. Using experimental designs Blader showed that when the nature of a procedure is clear, procedural elements shape perceived procedural justice. However, when procedures themselves were unclear, justice judgments were influenced by identification with the group and outcome favorability. In other words, nonfairness-related judgments became important in making justice judgments primarily when the justice of the situation was unclear.

Instrumental Views of the Psychology of Justice

One of the most important questions raised by the finding that justice matters is why people are motivated to act fairly. As I have noted, in theories of distributive justice the

answer is traditionally framed in social exchange terms (Thibaut and Kelley, 1959). People are [p. 352 ↓] viewed as being concerned with developing effective ways to exchange resources with others, both within particular situations and over time, since such cooperation is generally recognized as being to everyone's advantage. The development of principles of fairness occurs, from this perspective, because it aids in resource exchange.

Shared principles of fairness aid resource exchange because they indicate the distribution of resources which constitutes a reasonable exchange. Having such rules facilitates material exchanges, since there are clear rules for what each person deserves and each exchange does not have to begin with an effort to define reasonable exchange principles. Shared principles also facilitate the occurrence of exchanges since they allow people to alleviate their concerns that they are being disadvantaged in exchanges with others (acting like a "sucker"), or conversely that they are taking advantage of others. People can compare their outcomes to principles of justice to determine if what they are receiving in relationship to others is reasonable and appropriate. As a consequence, everyone can feel both secure and good about themselves during this process.

This argument suggests that having principles of distributive fairness is a precursor to effective cooperation and the ability to develop such shared principles may be a very fundamental aspect of people's social skills that has facilitated the evolution of humans into social beings who live in organized societies and cooperate. Recent research has supported this argument by demonstrating that animals that live in group settings, for example monkeys (Brosnan, 2006) and dogs (Range et al., 2009, recognize and act in accord with principles of distributive fairness. It is particularly striking that the members of both these species share with humans the willingness to forego rewards to defend principles of fairness.

Ironically, however, while these arguments support the idea that justice matters, they diminish the social psychological significance of justice findings because they suggest that people's motivation for caring about justice is their own material self-interest both immediately and over time. Social exchange models, such as that of Rusbult, argue that people in groups have a long-range view of their self-interest, often investing their efforts in groups in anticipation of long-term payoffs (e.g., Rusbult et al., 1991). If the

principles of distributive justice are accepted as instruments of coordination in the service of self-interest, they show a sophisticated ability on the part of both people and animals to develop coordinating rules and principles. But they do not suggest that people are motivated by intrinsic justice concerns when they act fairly.

An example of the implications of the social exchange argument is provided by discussions of the scope of justice. While some writers present the motivation to act justly as a core and universal human motivation (Lerner, 1980), others argue that it is bounded or limited in scope. The possibility of a scope of justice has important societal implications, since that scope can shift with events so that both individuals and the members of groups can become included, or excluded, from the scope of other's moral community. Once outside it, people are no longer accorded the presumption of treatment with dignity and respect for rights that group members in good standing assume they will receive (Nagata, 1993; Opotow, 1993).

Deutsch (1985) argues that people do not extend their concern about justice to all living things, or even to all people. Rather their concerns have a clear scope and outside of that scope people do not act in accord with the principles of distributive fairness. What defines this scope of justice? To Deutsch it is the domain of productive social exchange relationships. In other words, people follow principles of distributive justice with those with whom they see the potential of beneficial social exchanges, rather than feeling some type of intrinsic justice based motivation to act fairly. And those people or animals that are not viewed as candidates for productive social exchange are not treated with justice.

[p. 353 ↓]

As has been noted, research has demonstrated that in social settings issues of procedural justice are especially important in shaping people's thoughts, feelings, and behaviors. Does this focus on procedures, rather than outcomes, suggest a need for a shift in our understanding of the psychology of justice away from the model outlined above? The earliest psychological model of procedural justice is the control model presented by Thibaut and Walker (1975) and the control model is based upon the same ideas of social exchange that are presented in earlier discussions of distributive justice.

Thibaut and Walker assume that a person's goal when dealing with others is to achieve a distributively fair solution for themselves.

The arguments advanced by Thibaut and Walker in the context of procedural justice are similar to those outlined within the field of distributive justice – justice in the service of obtaining desired outcomes. A similar argument about the development of procedures is found in the work of Thibaut and Faucheux (1965) on the development of rules. Their argument is that rules develop to guide interactions when there is a risk that, without rules, a mutually beneficial exchange relationship will collapse. Hence, procedures/rules develop to facilitate productive resource exchanges.

In the case of procedures, Thibaut and Walker suggest that people prefer to keep control over their decisions, that is, to negotiate with others, and only turn to third-party procedures when necessary to protect productive social exchanges. Even then, people try to retain as much control as possible; for example, preferring mediation to arbitration. This instrumental view of procedures is illustrated in Thibaut and Faucheux's argument that procedures only develop when both parties have countervailing power. People are not viewed as intrinsically motivated to enact fair procedures, any more than they are intrinsically motivated to give others fair outcomes. They do so when they need to have rules or procedures to facilitate productive social exchange. In a situation in which relationships are vulnerable to disruption, people care about issues of justice (Barrett-Howard and Tyler, 1986).

Social Perspectives on the Psychology of Justice

Studies have not supported an instrumental view of justice. They support the argument that people want the opportunity to present their arguments to the decision maker, a procedural feature often labeled "voice." However, they have not supported the argument that people link voice to decision control and only value the opportunity to address the decision maker when they believe their arguments are shaping the outcome. Studies indicate that people value voice even when they do not believe that their voice leads to decision control (Lind et al., 1990; Tyler, 1987; Tyler et al., 1985).

These studies of voice suggest that having the opportunity for “voice” has interpersonal or “value-expressive” worth that was not linked to any influence over the decisions made (Tyler, 1987).

What factors were driving the influence of voice, even when it clearly could not affect the eventual outcome or decision? If an authority listens to people's arguments the authority is conferring interpersonal respect on that person because they are acknowledging their status in the group and their right as a group member to call upon the group about their needs and concerns. This argument is supported by the finding that people only value voice opportunities if they feel that the authority is “considering” their arguments (Tyler, 1987). This suggests that people focus on whether or not their concerns and needs in the situation are treated respectfully by the decision-maker, who takes them seriously by listening and considering what they have to say, independently of whether or not the course of action the authority recommends reflects their desired course of action.

These findings lead to the group-value model (Lind and Tyler, 1988), which focuses [p. 354 ↓] on the antecedents of judgments of procedural justice. The group-value model argues that noninstrumental factors influence procedural justice judgments, a prediction confirmed both by the findings of noninstrumental voice effects already noted (Lind et al., 1990; Tyler, 1987), and by demonstrations that people care more about a broader set of issues of procedural justice when dealing with members of their own groups (Tyler, 1999); issues including how they are treated, and whether their rights are respected. These noninstrumental issues are important because they communicate information to people about their status within groups; that is, because they carry an important social message (Lind and Tyler, 1988). This suggestion led Lind and Tyler to draw upon the ideas of social identity theory and argue that people value voice because it shapes their identity and provides information about their status in groups.

Of course, it is important to note that like prior models the group value model also argues for a scope of justice. In this case that scope is defined by the range of people or groups that are relevant to people's definitions of their status, that is, to the range of their group. For example, people are less concerned about justice when they are dealing with outsiders (Smith et al., 1998). Further, those people who are less concerned about issues of their status – that is, the quality of their social connections

— are generally less influenced by information about justice (DeCremer and Tyler, 2005; Tyler and Lind, 1992). A typical American, for example, is likely to be relatively indifferent to their status in Japanese society, so they are unaffected by variations in treatment by Japanese authorities since that treatment does not communicate information about their status in their own group.

Recent studies demonstrate that it is possible to prime people so that they are focused upon instrumental or relational issues. As would be predicted, instrumental priming leads people to focus upon the anticipated outcomes of third-party decisions, reacting to what they receive. Relational priming, on the other hand, leads people to focus upon the fairness of decision-making procedures (Stahl et al., 2008).

The Distinction between Justice and Injustice

Justice theories argue that a variety of types of reaction follow from justice judgments, including most importantly behaviors such as retaliation. Distributive justice research focuses primarily upon anger and negative behaviors, that is, upon reactions to injustice, since this research is rooted in the literature on relative deprivation, a literature whose origins lie in efforts to understand and explain riots and rebellion (Gurr, 1970). This focus on negative attitudes and behaviors is reflected in later efforts to understand distributive influences on pay dissatisfaction, employee theft, sabotage, turnover, and resistance to third-party decisions (Tyler and Smith, 1998).

The relational model also focuses upon reactions to negative events in the form of poor outcomes and predicts that procedural justice will influence reactions to the authorities who deliver them (Tyler, 2006a, 2006b). Its predominate focus is on decision acceptance. Recent research on procedural justice has increasingly focused on more prosocial outcomes, such as how to build trust, encourage responsibility and obligation, generate intrinsic motivation and creativity, and stimulate voluntary cooperation with others (Tyler and Blader, 2000). Similarly, there has been increasing attention to exploring when justice motivations encourage people to provide resources to the disadvantaged (Montada, 1995). Interestingly, this shift is consistent with a shift that

has been taking place within psychological research more generally (Snyder and Lopez, 2002).

This new focus of justice research is addressed by the group engagement model, which discusses the antecedents of cooperative behavior in groups, organizations, and [p. 355 ↓] societies (Tyler and Blader, 2000, 2003). The argument of the group engagement model is that justice theories provide a basis for understanding people's general relationship to groups. That includes both people's negative reactions to injustice and the ability of experiencing justice to promote engagement and cooperation. Society, after all, does not just want people not to riot or destroy. It also wants them to be happy, creative, and productive.

How does the group engagement model expand earlier models? First, the objective of the model is to identify and examine the antecedents of attitudes, values, and cooperative behavior in groups. Hence, the group engagement model broadens the focus of justice studies and its predecessor models of justice by positing a general model of the relationship between people and groups. In trying to understand the precursors of people's engagement in their groups, it identifies and examines a much broader set of variables – and dynamics between those variables – than earlier justice models. Rather than focusing simply on what shapes views about justice the model is concerned with the role of justice in social systems.

People have considerable discretion about the degree to which they invest themselves in their groups by working on behalf of the group. To examine this issue, the group engagement model distinguishes between two classes of cooperative behavior: mandatory and discretionary. Mandatory cooperation is behavior that is stipulated by the group, while discretionary cooperation originates with the group member. The model argues that each of these forms of cooperation is differently motivated. Of the two types of cooperative behaviors, mandatory behaviors are more strongly affected by incentives and sanctions, since they are behaviors required by the group and thus the group specifically structures incentives and sanctions to encourage these behaviors. Discretionary behaviors are more strongly under the influence of people's internal motivations (their attitudes and values), since they are behaviors that originate with the individual.

Both attitudes and values are important because they lead people to be internally motivated to engage in and cooperate with the group. To the degree that people are internally motivated, they engage in cooperative behaviors for personal reasons, and they do not need to receive incentives (rewards) or to face the risk of sanctions (punishments) to encourage their group-related behaviors. This benefits groups, which are then free to deploy their assets in other ways that benefit the group.

The Influence of Identity and Resource Motivations on Engagement in Groups

The group engagement model argues that the central reason that people engage themselves in groups is because they use the feedback they receive from those groups to create and maintain their identities. In other words, the group engagement model hypothesizes that of the two types of motivations, it is the development and maintenance of a favorable identity that most strongly influences cooperation. The model predicts that people's willingness to cooperate with their group – especially cooperation that is discretionary in nature – flows from the identity information people receive from the group.

The core argument of the group engagement model is that people want a favorable identity, and use group membership as one source of identity-relevant information. That identity information, in turn, is hypothesized to emanate from evaluations of the justice experienced in the group. This includes judgments of procedural and distributive justice, as well as evaluations of outcome favorability. This suggests that identity evaluations and concerns mediate the relationship between justice judgments and group engagement. This is the *identity mediation hypothesis*.

Why might this be so? Using social identity theory as a framework the model argues that an important function of groups is to provide people with a way of constructing [p. 356 ↓] a social identity. It is widely recognized that groups shape people's definitions of themselves and their feelings of wellbeing and self-worth (Hogg and Abrams, 1988; Sedikides and Brewer, 2001). In particular, group memberships shape people's conceptions of their social selves – the aspect of the self that is formed through

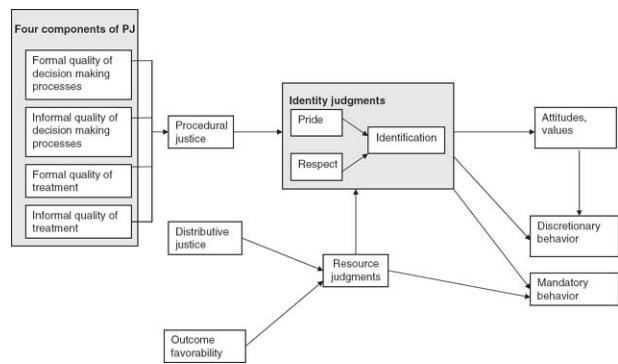
identification with groups. The groups that people belong to help to define who they are and help people to evaluate their status. Part of this process involves linking views about self-worth and self-esteem to the status of group memberships.

Thus, to some degree people's sense of their own worth is linked to the groups to which they belong. Several new ideas and hypotheses flow from the group engagement model. It predicts that identity judgments will be the primary factors shaping attitudes, values, and cooperative behaviors in groups. Second, it predicts that resource judgments will most strongly influence attitudes, values, and discretionary cooperative behaviors in groups through their indirect influence on identity judgments, rather than directly. Third, it predicts that the primary antecedent of identity judgments will be judgments about the procedural justice of the group. Fourth, it predicts that status judgments about pride and respect will shape identification with the group. Each of these predictions is reflected in [Figure 43.1](#).

The focus of the group value model of procedural justice; the relational model of authority; and the group value model of engagement is on the psychology of justice – that is, on why justice matters. All three models argue that procedural justice, the form of justice that is central to people's connection to groups, is linked to issues of status and identity. Because of this centrality of issues of identity, this discussion of justice involves attention to identity its role in people's behavior in groups. The group engagement model argues that identity plays an important role in people's relationship to their group. It focuses on what people *get from groups* in the form of acknowledgement and recognition of their identities.

As has been noted, the identity-based model of cooperation can be contrasted to a resource-based model of cooperation. Social psychologists have long recognized that [p. 357 ↓] people interact with others to exchange material resources. These material resources can vary widely – from things such as food to money – but regardless they share the characteristic of being material resources that people obtain through their cooperation with others (Thibaut and Kelley, 1959). Guided by social exchange models, many social psychological discussions of people's relationships to groups have argued that this exchange of material resources is the fundamental reason that people engage in groups.

Figure 43.1 Group engagement model



A broadened model of social exchange perspective is the basis of several more recent resource-based models, including the investment model, which focuses on exit and loyalty to groups (Rusbult and Van Lange, 1996) and realistic group conflict theory (Levine and Campbell, 1972). These resource-based perspectives predict that people's level of cooperation with a group will be shaped by the level of the material resources that they receive from that group and/or the sanctioning risks they face within the group. Thus, the willingness to voluntarily cooperate with the group by doing things that help the group flows from assessments of the desirability of the resources that are gained or lost by association with the group. In addition, loyalty to the group will also be shaped by the level of resources people are obtaining, relative to what they might obtain in another group. More recent models of social exchange also recognize that concerns about material gain and loss can be transformed in social interactions to reflect broader social values (see Van Lange et al., 2007) and, in that sense, are similar to the group engagement model in suggesting that nonmaterial issues play an important role in shaping links between people and groups.

The group engagement model proposes that the identity model prevails over the resource model in predicting engagement and cooperation. It argues that resource judgments do not directly shape engagement. This is not to say that the group engagement model argues that resource judgments have no influence on engagement. Instead, the model hypothesizes that resource judgments indirectly influence most forms of engagement by shaping identity. That is, to some degree, people evaluate their identity and status in a particular group by the level of the resources that they are

receiving from that group. To the extent that having more resources in a group leads people to feel better about their status in the group, they will engage themselves more in that group.

The group engagement model argues that this is not the case and that such material rewards primarily influence engagement indirectly, by influencing status. The key argument of the group engagement model is that people's level of cooperation with groups is primarily shaped by the extent to which they identify with those groups. Cooperation is driven, in other words, by the motivation to create and maintain a favorable identity, and that identity flows from justice.

It seems counterintuitive to many people to argue that resources are not the primary factor that directly shapes engagement. Certainly, people can think of many examples from their everyday lives that seem to suggest a resource-based linkage with engagement. The seeming importance of resource concerns is also supported by some research findings. This may reflect evidence of the indirect connection between resource judgments and engagement of the type we have already outlined. If, as the group engagement model argues, resource judgments indirectly influence engagement, then studies that do not measure identity judgments will find a connection between resource judgments and engagement. However, the group engagement model suggests that in a fully specified model, which includes both resource and identity judgments, the spurious connection between resource judgments and engagement will disappear (except for that between resources and mandatory cooperation), while a mediated connection remains.

How does the group matter? The group engagement model also considers how the policies and practices of the group shape [p. 358 ↓] identity-based and resource-based judgments. It is this aspect of the group engagement model that directly addresses issues of justice. The group engagement model argues that people are most strongly influenced by one aspect of the policies and practices of their group – the fairness of the group's procedures. This argument builds on the pervasive finding that procedural justice judgments have a strong and widespread influence on people's thoughts, feelings, and behaviors in group contexts (Lind and Tyler, 1988; Tyler and Smith, 1998). In addition, judgments about outcome fairness and outcome favorability both shape identity. Hence, both forms of justice shape identity.

As already noted, it is also possible to conceptualize a relationship between the person and the group that is centered around the exchange of resources. If this were the key motivation that shapes people's engagement in groups, it would be expected that the element of group policies and practices that would most shape their engagement is their estimate of the degree to which the rules and policies of the group provide them with desirable resources. These resource judgments, in turn, influence engagement in the group. In the case of either outcome fairness or outcome favorability, it is the concern over the outcomes that are being received from the group that would be driving engagement in groups. The group engagement model argues that such outcome based judgments influence identity. Hence, people are more likely to identify with groups that deliver desired resources.

Are the hypotheses of the group engagement model valid? Tyler and Blader (2000) tested the model using survey data from employees. Using causal modeling, they tested several of the key hypotheses of the group engagement model and found support for all of them (see Tyler and Blader, 2000: 196). First, they found that identity judgments shaped attitudes, values, and cooperative behaviors. Consistent with the predictions of the model, they found a greater influence of identity judgments on discretionary, as compared to mandatory, behavior. Second, resource judgments are found to influence attitudes, values, and discretionary cooperative behaviors indirectly, through identity judgments, but not directly. Third, procedural justice judgments are found to be the primary antecedent of identity judgments (Tyler and Blader, 2000: 136). Tyler (in press) confirms these findings in both work organizations and communities using panel data, while Blader and Tyler (in press) do so using independently derived indices of cooperative behavior.

A Broader Identity Framework

The underlying argument of the group-value model of procedural justice, the relational model of authority, and the group engagement model of cooperation is that identity plays a central mediating role in shaping people's reaction to groups, organizations and societies. This argument was originally put forward in the context of social identification theories (Hogg and Abrams, 1988). In developing this model identification, the merger of self with the group, was distinguished from pride in the group, which is linked to

group status, and respect from the group, which is the consequence of status in the group. This aspect of self, as opposed to the personal self (unique individual traits), or the relational self (the self defined by dyadic relationships) is the collective self and is linked to group memberships (Sedikides and Brewer, 2001). More recently DeCremer and Tyler (2005) have broadened the framework within which identity is studied within the procedural justice literature. In particular, they demonstrate that whenever people link their identities to a group the justice that they experience in the framework of that group has a stronger influence upon their sense of self. This supports the argument that justice is linked to identity and, therefore, justice matters more when identity is more relevant.

[p. 359 ↓]

Conclusion

The findings of justice research are important for several reasons. First, they contribute to the demonstration that people's thoughts, feelings, and behaviors are determined by their internally held values concerning what is just or fair. These values play an important role in making social life possible because they provide a basis for cooperation among people in groups, organizations, and societies. And, as the literature on social justice makes clear, they provide an important confirmation that the social ties between people are central to their actions in social settings. People in social settings do not act simply as self-interested actors, pursuing individual or group gains and losses. Rather their feelings, thoughts and behaviors are shaped by their judgments concerning that is appropriate and fair. The demonstration that people are value-based actors provides a clear demonstration of the centrality of social motivations to people's actions in groups, communities, organizations and societies.

It is not obvious that people's engagement in groups would be the result of procedural justice judgments. People could potentially consider a wide variety of aspects of their relationship to their group when they are evaluating the degree to which they want to engage themselves in a group. One thing that we might expect people to consider is reward level – that is, people might consider their salaries, the number of resources they are given to manage, and/or the size of their office, their car, or their home as

key inputs into their judgments about how much to engage themselves in a group. Or, at least, they might consider outcome fairness, as suggested by Thibaut and Walker (1975).

Because an outcome focus is intuitively obvious, the finding that procedural justice is so central to people's thinking is striking. It is especially striking because, of the procedural elements considered, questions of interpersonal treatment consistently emerge as important. In other words, people's focus is upon those aspects of their experience that communicate messages about status, rather than upon those more directly related to issues of decision making. This supports the argument that it is status issues that define people's relationship to groups, and procedural justice that provides information about status.

Overall, the literature on justice contributes to a social vision of the person on several levels: first, because people care about justice, a socially constructed idea, and view it as the core element of social groups; second, because people think of justice in very relational terms; and third, because studies of how justice influences people's behavior suggests that the key connection between people and groups, communities, organizations and societies is rooted in their concerns about self and identity. In all of these ways, people show themselves to be fundamentally social animals, concerned about their relationships with others.

Tom R.Tyler

References

Barrett-Howard, E. and Tyler, T.R. Procedural justice as a criterion in allocation decisions . Journal of Personality and Social Psychology , 50, 296–304. (1986) <http://dx.doi.org/10.1037/0022-3514.50.2.296>

Bies, R.J. and Moag, J.S. (1986) Interactional justice . In R.J. Lewicki, ed. , B.M. Sheppard, ed. , and M.H. Bazerman (eds), Research on Negotiations in Organizations , 1, 43–55. Greenwich, CT: JAI.

Blader, S.L. What determines people's fairness judgments? Identification and outcomes influence procedural fairness evaluations under uncertainty . Journal of Experimental Social Psychology , 43, 986–994. (2007) <http://dx.doi.org/10.1016/j.jesp.2006.10.022>

Blader, S.L. and Tyler, T.R. A four component model of procedural justice . Personality and Social Psychology Bulletin , 29, 747–758. (2003) <http://dx.doi.org/10.1177/0146167203029006007>

Blader, S.L. and Tyler, T.R. Testing and extending the group engagement model: Linkages between social identity, procedural justice, economic outcomes and extrarole behavior . Journal of Applied Psychology . 94, 445–464. (2009) <http://dx.doi.org/10.1037/a0013935>

Brosnan, S.F. Nonhuman species' reactions to inequity and their implications for fairness . Social Justice Research , 19, 153–185. (2006) <http://dx.doi.org/10.1007/s11211-006-0002-z>

Chen, E. and Tyler, T.R. (2001) Cloaking power: Legitimizing myths and the psychology of the advantaged . In J.A. Bargh, ed. and A.Y. Lee-Chai (eds), *The Use and Abuse of Power* , pp. 241–261. Philadelphia: Psychology Press.

Darley, J.M. and Pittman, T.S. The psychology of compensation and retributive justice . Personality and Social Psychology Review , 7, 324–336. (2003) <http://dx.doi.org/10.1037/0022-3514.78.5.853>

DeCremer, D. and Tyler, T.R. Managing group behavior: The interplay between procedural justice, sense of self, and cooperation . Advances in Experimental Social Psychology , 37, 151–218. (2005) <http://dx.doi.org/10.1016/S0065-2601%2805%2937003-1>

Deutsch, M. Equity, equality, and need: What determines which value will be used as the basis for distributive justice? Journal of Social Issues , 31, 137–149. (1975) <http://dx.doi.org/10.1111/j.1540-4560.1975.tb01000.x>

Deutsch, M. (1985) *Distributive Justice* . New Haven: Yale University PRes.

Gold, M. (1999) *The Complete Social Scientist: A Kurt Lewin Reader*. Washington, DC: American Psychological Association.

Gurr, T. (1970) *Why Men Rebel*. Princeton: Princeton University Press.

Handgraaf, M.J.J., Van Dijk, E., Wilke, H.A.M. and Vermunt, R. Evaluability of outcomes in ultimatum bargaining. *Organizational Behavior and Human Decision Processes*, 95, 97–106. (2004) <http://dx.doi.org/10.1016/j.obhd.2004.06.005>

Hogg, M.A. and Abrams, D. (1988) *Social Identifications*. London: Routledge.

Huo, Y.J. Justice and the regulation of social relations: When and why do group members deny claims to social goods? *British Journal of Social Psychology*, 41, 535–562. (2002) <http://dx.doi.org/10.1348/014466602321149876>

Kelley, H.H., Holmes, J.G., Kerr, N.L., Reis, H.T., Rusbult, C.E. and Van Lange, P.A.M. (2003) *An Atlas of Interpersonal Situations*. Cambridge: Cambridge University Press.

Kelley, H.H. and Thibaut, J.W. (1978) *Interpersonal Relations: A Theory of Interdependence*. New York: Wiley.

Lerner, M.J. (1980) *The Belief in a Just World: A Fundamental Delusion*. New York: Plenum.

Leventhal, G.S. (1980) What should be done with equity theory? New approaches to the study of fairness in social relationships. In K. Gergen, ed. , M. Greenberg, ed. , and R. Willis (eds), *Social Exchange*, pp. 27–55. New York: Plenum. http://dx.doi.org/10.1007/978-1-4613-3087-5_2

Levine, R.A. and Campbell, D.T. (1972) *Ethnocentrism: Theories of Conflict, Ethnic Attitudes, and Group Behavior*. New York: Wiley.

Lind, E.A., Kanfer, R. and Earley, C. Voice, control, and procedural justice. *Journal of Personality and Social Psychology*, 59, 952–959. (1990) <http://dx.doi.org/10.1037/0022-3514.59.5.952>

Lind, E.A. and Tyler, T.R. (1988) *The Social Psychology of Procedural Justice* . New York: Plenum.

Messick, D.M., Bloom, S., Boldizar, J.P. and Samuelson, C.D. Why we are fairer than others . *Journal of Experimental Social Psychology* , 21, 389–399. (1985) <http://dx.doi.org/10.1016/0022-1031%2885%2990031-9>

Messick, D. and Sentis, K. Estimating social and nonsocial utility functions from ordinal data . *European Journal of Social Psychology* , 15, 389–399. (1985) <http://dx.doi.org/10.1002/ejsp.2420150403>

Mikula, G., Petri, B. and Tanzer, N. What people regard as unjust . *European Journal of Social Psychology* , 22, 133–149. (1990) <http://dx.doi.org/10.1002/ejsp.2420200205>

Montada, L. Applying social psychology: The case of redistributions in united Germany . *Social Justice Research* , 8, 73–90. (1995) <http://dx.doi.org/10.1007/BF02334827>

Nagata, D. (1993) *Legacy of Injustice* . New York: Plenum.

Opotow, S. Animals and the scope of justice . *Journal of Social Issues* , 49, 71–86. (1993) <http://dx.doi.org/10.1111/j.1540-4560.1993.tb00909.x>

Range, F., Horn, L., Viranyi, Z. and Huber, L. The absence of reward induces inequity aversion in dogs . *Proceedings of the National Academy of Science* , 106, 340–345. (2009) <http://dx.doi.org/10.1073/pnas.0810957105>

Rawls, J. (1971) *A Theory of Justice* . Cambridge: Harvard University Press.

Ross, M. and Sicoly, F. Egocentric biases in availability and attribution . *Journal of Personality and Social Psychology* , 37, 322–336. (1979) <http://dx.doi.org/10.1037/0022-3514.37.3.322>

Rusbult, C. and Van Lange, P. (1996) Interdependence processes . In E.T. Higgins, ed. and A.W. Kruglanski (eds), *Social Psychology* , pp. 564–596. New York: Guilford Press.

Rusbult, C.E., Verette, J., Whitney, G.A., Slovik, L.F. and Lipkus, I. Accommodation processes in close relationships: Theory and preliminary empirical evidence . Journal of Personality and Social Psychology , 60, 53–78. (1991) <http://dx.doi.org/10.1037/0022-3514.60.1.53>

Sedikides, C. and Brewer, M.B. (2001) Individual Self, Relational Self, and Collective Self . Philadelphia: Psychology Press.

Smith, H.J., Tyler, T.R., Huo, Y.J., Ortiz, D.J. and Lind, E.A. The self-relevant implications of the group-value model: Group membership, self-worth, and procedural justice . Journal of Experimental Social Psychology , 34, 470–493. (1998) <http://dx.doi.org/10.1006/jesp.1998.1360>

Snyder, C.R. and Lopez, S.J. (2002) Handbook of Positive Psychology . New York: Oxford University Press.

Stahl, T., Vermunt, R. and Ellemers, N. For love or money? How activation of relational versus instrumental concerns affects reactions to decision-making procedures . Journal of Experimental Social Psychology , 44, 80–94. (2008) <http://dx.doi.org/10.1016/j.jesp.2007.02.004>

Tajfel, H. and Turner, J. (1979) An integrative theory of intergroup conflict . In W.G. Austin, ed. and S. Worchel (eds), The Social Psychology of Intergroup Relations , pp. 33–47. Monterey, CA: Brooks/Cole.

Thibaut, J. and Faucheuex, C. The development of contractual norms in a bargaining situation under two types of stress . Journal of Experimental Social Psychology , 1, 89–102. (1965) <http://dx.doi.org/10.1016/0022-1031%2865%2990038-7>

Thibaut, J. and Kelley, H.H. (1959) The Social Psychology of Groups . New York: Wiley.

Thibaut, J. and Walker, L. (1975) Procedural Justice . Hillsdale, NJ: Erlbaum.

Thompson, L. and Loewenstein, G. Egocentric interpretations of fairness and interpersonal conflict . Organizational Behavior and Human Decision Processes , 51, 176–197. (1992) <http://dx.doi.org/10.1016/0749-5978%2892%2990010-5>

Tyler, T.R. Conditions leading to value expressive effects in judgments of procedural justice: A test of four models . *Journal of Personality and Social Psychology* , 52, 333–344. (1987) <http://dx.doi.org/10.1037/0022-3514.52.2.333>

Tyler, T.R. What is procedural justice?: Criteria used by citizens to assess the fairness of legal procedures . *Law and Society Review* , 22, 103–135. (1988) <http://dx.doi.org/10.2307/3053563>

Tyler, T.R. The psychology of procedural justice: A test of the group value model . *Journal of Personality and Social Psychology* , 57, 830–838. (1989) <http://dx.doi.org/10.1037/0022-3514.57.5.830>

Tyler, T.R. Psychological models of the justice motive . *Journal of Personality and Social Psychology* , 67, 850–863. (1994) <http://dx.doi.org/10.1037/0022-3514.67.5.850>

Tyler, T.R. Why people cooperate with organizations: An identity-based perspective . *Research in Organizational Behavior* , 21, 201–246. (1999)

Tyler, T.R. Social justice: Outcome and procedure . *International Journal of Psychology* , 35, 117–125. (2000) <http://dx.doi.org/10.1080/002075900399411>

Tyler, T.R. (2006) *Why People Obey the Law* . Princeton: Princeton University Press.

Tyler, T.R. and Bies, R. (1990) Interpersonal aspects of procedural justice . In J.S. Carroll (ed.), *Applied Social Psychology in Business Settings* , pp. 77–98. Hillsdale, NJ: Erlbaum.

Tyler, T.R. and Blader, S.L. (2000) Cooperation in Groups: Procedural Justice, Social Identity, and Behavioral Engagement . Philadelphia: Psychology Press.

Tyler, T.R. and Blader, S.L. Procedural justice, social identity, and cooperative behavior . *Personality and Social Psychology Review* , 7, 349–361. (2003) http://dx.doi.org/10.1207/S15327957PSPR0704_07

Tyler, T.R., Boeckmann, R., Smith, H.J. and Huo, Y.J. (1997) *Social Justice in a Diverse Society* . Denver, CO: Westview.

Tyler, T.R. and Caine, A. The role of distributional and procedural fairness in the endorsement of formal leaders . *Journal of Personality and Social Psychology* , 41, 642–655. (1981) <http://dx.doi.org/10.1037/0022-3514.41.4.642>

Tyler, T.R., Degoey, P. and Smith, H. Understanding why the justice of group procedures matters: A test of the psychological dynamics of the group-value model . *Journal of Personality and Social Psychology* , 70, 913–930. (1996) <http://dx.doi.org/10.1037/0022-3514.70.5.913>

Tyler, T.R. and Lind, E.A. A relational model of authority in groups . *Advances in Experimental Social Psychology* , 25, 115–191. (1992) <http://dx.doi.org/10.1016/S0065-2601%2808%2960283-X>

Tyler, T.R., Rasinski, K. and Spodick, N. The influence of voice on satisfaction with leaders: Exploring the meaning of process control . *Journal of Personality and Social Psychology* , 48, 72–81. (1985) <http://dx.doi.org/10.1037/0022-3514.48.1.72>

Tyler, T.R. and Smith, H.J. (1998) Social justice and social movements . In D. Gilbert, ed. , S. Fiske, ed. and G. Lindzey (eds), *Handbook of Social Psychology*, Vol. 2 , 4th Edition, pp. 595–629. New York: McGraw-Hill.

Van Lange, P.A.M., De Cremer, D., Van Dijk, E. and Van Vugt, M. (2007) Self-interest and beyond . In A.W. Kruglanski, ed. and E.T. Higgins (eds), *Social Psychology: Handbook of Basic Principles* , 2nd Edition. New York: Guilford Press.

Vidmar, N. and Miller, D.T. The social psychology of punishment . *Law and Society Review* , 14, 565–602. (1980) <http://dx.doi.org/10.2307/3053193>

Wakslak, C.J., Jost, J.T., Tyler, T.R. and Chen, E. System justification and the alleviation of emotional distress . *Psychological Science* , 18, 267–274. (2007) <http://dx.doi.org/10.1037/h0033967>

Walster, E., Walster, G.W. and Berscheid, E. (1978) *Equity: Theory and Research* . Boston: Allyn and Bacon.

<http://dx.doi.org/10.4135/9781446249222.n43>

Handbook of Theories of Social Psychology

Minority Influence Theory

Contributors: Charlan Jeane Nemeth

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Minority Influence Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n44>

Print pages: 362-379

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n44>

[p. 362 ↓]

Chapter 44: Minority Influence Theory

Abstract

The study of minority influence began as a reaction to the portrayal of influence as the province of status and numbers and from a realisation that minorities need not just be passive recipients of influence but can actively persuade. From these beginnings, a considerable body of research, including ours, has investigated how minority views prevail. In the decades that followed, we concentrated, not so much on persuasion or attitude change, but rather on the value of minority views for the stimulation of divergent thinking. Dissent, as has been repeatedly documented, ‘opens’ the mind. People search for information, consider more options and, on balance, make better decisions and are more creative. Dissenters, rather than rogues or obstacles, provide value: they liberate people to say what they believe and they stimulate divergent and creative thought even *when they are wrong*. The implications for group decision making, whether in juries or companies, have been considerable and there is increasing interest in research and in practice for the value of authentic dissent in teams and in creating ‘cultures’ of innovation.

Introduction

My lifelong interest has been the study of influence – and in particular, influence by those who hold minority opinions. Initially, we concentrated on how minorities ‘win’ or persuade others to their position. Stimulated by observations on juries, we recognised the potential value of dissent, not for the truth that it holds or for its ability to persuade, but rather for the thought that it stimulates. Repeatedly we found that dissent stimulates thought that is more enquiring, more divergent and more creative. By contrast, majority views stimulate convergent thinking. People focus on the issue from the perspective of

the majority and narrow the range of considerations, often convincing themselves of the majority position.

This work has had influence on the dialogue within social psychology but also on the law as well as corporate cultures in organisations. Dissent has come to be seen as having value and not simply as an 'obstacle'. There is serious consideration of the importance of protecting dissent in juries via procedural rules such as the requirement of unanimity. And, in organisations, the assumptions regarding the value of cohesion and homogeneity have been complicated with a willingness to recognise the importance of diversity and dissent for innovation.

[p. 363 ↓]

Background

As a graduate student in the mid 1960s, I had the good fortune of studying with Henri Tajfel and, through him, meeting Serge Moscovici (in England and France respectively) and, by virtue of their influence, I recommitted to the reasons why I decided to study social psychology. As an undergraduate math major, it was a single lecture on brainwashing that motivated me to pursue graduate studies in social psychology; it was the simple but compelling observation that people are powerful sources of influence. They can educate, inspire and strengthen us or they can diminish and weaken us as human beings. Though graduate school proved to be disappointing with careful, highly choreographed and often single variable studies, it was a 'gap' year in Oxford that renewed my interests. Perhaps it took two Eastern European Jews who went through World War II to teach me that we could and should study the combat between ideas and groups, that there were exemplars of courage and cowardice brought about by social conditions and the influence of others.

It was my time as a visiting professor in Bristol with Henri and in Paris with Serge, the year after my doctorate, that was to have the most influence on my thinking and subsequent professional career. Henri was passionate about categorisation and the importance of ingroups and outgroups. Serge was rethinking the flow of influence.

Social psychology at that time portrayed influence as flowing from the strong to the weak. We learned the value of status, confident styles and numbers. High status individuals (or the perception of higher status) influenced those of lower status (Berger et al., 1977; Hovland et al., 1953). It was the white, tall, wealthy, attractive male who studied at elitist institutions and/or who had position and title who influenced those of us who had few of those attributes or demographics. People who spoke loudly, quickly and with authority exercised influence (Giles et al., 1979). And we certainly learned about the power of majorities. Faced with a disagreeing majority, even one that was wrong, people abdicated information from their own senses and became subject to doubt (Asch, 1956; Kiesler and Kiesler, 1970). The individual or minority of individuals were passive agents. Their only choice seemed to be one of independence or conformity to the majority view.

The power of these elements was true then and it is true now. However, not being in the categories of status nor numbers in psychology (who were predominantly male) nor being particularly loud or authoritative, I can remember feeling relatively powerless in the influence world. Much of the research, including that later embodied in social impact theory (Latané, 1981) conveyed the imagery of a large truck that pummelled people into acquiescence. Perhaps Tajfel and Moscovici and the students who were drawn to them in those days identified with being in a minority and thus found appeal in a conception of an active minority view, but it was also the world in which we found ourselves. During that year, Serge and I spent a lot of time discussing minority views and their power. He would often tell me that American social psychology did not portray the world he saw. The emphasis on winning friends and influencing people, reciprocating favours and continually resolving conflicts was in contrast to the experiences of his generation. Being 1970, it was also in contrast to both his and my then recent experience. He went through the May 1968 uprising which had lasting effects on the university and French society in general.

I, as a brand new assistant professor at the University of Chicago (1968), I saw in person a vigorous conflict of ideas about the Vietnam War. The antiwar minority did not conform or remain silently independent. They were vocal, persistent, and aggressive – and there were violent consequences in response. We witnessed the beatings, the ‘lessons’ taught to faculty who supported those protests but we also saw social change and an evolution of public opinion that was dramatic.

[p. 364 ↓]

Shortly before, Serge along with Claude Faucheux developed a theory questioning the one-sided conceptualisation of influence (Faucheux and Moscovici, 1967; Moscovici and Faucheux, 1972). Recognising that social change cannot be understood without attention to an initial minority view, they set out to demonstrate that minority views can prevail. As we had all witnessed, individuals and minorities were not simply passive agents who either resisted or conformed to majority judgments; they could in fact exercise influence themselves. Furthermore, behavioural style is important and, in particular, consistency over time is key to exerting influence. Not having the numbers or status to influence at the outset, a member of a minority needed to argue his/her position effectively. It is the orchestration and patterning of verbal and nonverbal behaviour that fosters influence (Moscovici and Nemeth, 1974). And in 1969, an experiment was able to demonstrate just such an effect (Moscovici et al., 1969). That study and its findings have been replicated numerous times and provide, in my judgment, some of the best insights into why and how minorities can prevail. Thus it is worth describing in some detail.

In that study (which is almost a mirror image of a typical conformity experiment), four naïve individuals and two confederates were shown a series of blue slides and were asked to judge the colour and the brightness of those slides. A control group consisting of only naïve individuals had no problem; they repeatedly called the blue slides 'blue'. In the experimental conditions, the two confederates offered a differing view. In the 'consistent' condition, they repeatedly called the blue slides 'green'. In other words, they said 'green' on every trial. In the 'inconsistent' condition, they called the slides 'green' on two-thirds of the trials and called them 'blue' on one-third of the trials.

Most people – and certainly most students in my courses – hypothesise that the inconsistent condition would exert more influence. After all, they are correct on one-third of the trials and, further, they agree with the majority on those trials and are thus likely to be better liked. The findings were the reverse. In keeping with the theory emphasising the importance of consistency, the findings showed that the consistent minority persuaded the majority to say 'green' on nearly nine per cent of trials whereas the inconsistent minority exerted no influence, not differing statistically from the control group.

Perhaps more importantly, the findings from this study showed that influence was greater at a private or latent level than evident at a public level. After the experimental procedure, individuals were asked to sort a series of blue/green chips into the categories of either 'green' or 'blue'. If you think of the coloured chips you could find in a paint store, you can order the blue/green chips along an actual physical dimension – from 'blue' to 'green'. You can mix up these chips but, at a point along that physical dimension, individuals or a control group call those to the 'blue end' of that point 'blue' and to the 'green end' of that point 'green'. Experimental subjects in the consistent condition shifted that point. They called the chips 'green' when a control group would call them 'blue'. These two findings have stood the test of time.

As I reflect back on conversations in those early years, Serge was very clear about the importance of consistency – and the perils of compromise. He had a rather psychoanalytic view of influence. You had to combat the resistances. There was a war of ideas; there was conflict not easily resolved; and above all, there was clarity, even with exaggeration and omission.

Serge behaved as he theorised and his lessons to me were personal as well as professional. He always put my feet to the fire. Whenever I would lapse in the graduate student habit of citing other people to legitimate my point, he would ask what I really believed followed by 'you must write *that*'. He convinced me that consistency and confidence were imperative. Know what you know and speak what you believe. But in those observations, it also became clear that influence involved 'style over time', that there was an [p. 365 ↓] orchestration of the verbal and nonverbal ways in which people interacted and persuaded. This stimulated an early study showing that consistency could be maintained without the confrontation of repetition and increasing the conflict, provided that changes in position were due to changes in the stimuli or 'facts' (Nemeth et al., 1974). Above all, we came to recognise that there was a subtlety about influence, a public face that often masked what people really thought.

Behavioural Style

The early study demonstrating the importance of consistency over time has been repeatedly corroborated and developed (Moscovici and Lage, 1976; Moscovici

and Nemeth, 1974; Mugny, 1982; Nemeth et al., 1974, 1990). For the minority to be ‘persuasive’, they need to be seen as having a position in which they believe. Minimally this means consistency of the position, not necessarily repetition. Mugny has further pointed to the subtlety of consistency by making the distinction between flexibility and rigidity, the former but not the latter being conducive to persuasion. Additionally, numerous studies have confirmed that minority influence is aided by confident behavioural styles; for example, taking the head seat (Moscovici and Lage, 1976; Nemeth and Wachtler, 1974) or even by a perception of confidence which might occur by virtue of simply persisting as a minority voice (Nemeth et al., 1977). This area, in my judgment, is both fascinating and worthy of further development though the subtleties involved in the timing and choreography of behavioural style are challenging for researchers.

The Private or Latent Nature of Influence

Perhaps the more important development from that early study is the fact that minority views exercise their influence at a private or latent level rather than in public. Even in our early discussions, it was clear to us that influence was deeper and more profound than public adoption of a position. People may not publicly agree with the minority position but they may agree privately, later or in a different form (David and Turner, 2001; Mugny et al., 1995; Nemeth and Wachtler, 1974). It may appear on indirect judgments or be generalised to other subjects (Forgas and Williams, 2001; Maass et al., 1987; Mugny, 1982).

Part of the reason for the ‘latent’ influence by minorities is that people do not want to publicly move to a minority view since they fear the likely ridicule and rejection, consequences that have been documented since the early work by Schachter (1951).

In one of our own studies (Nemeth and Wachtler, 1974), we saw very clearly how adamantly subjects would resist any public movement to a minority position. The setting was a mock jury involving a personal injury case. The case was hypothetical, but the anger was so evident that subjects were pounding their fists on the table next to the confederate’s face (the one who argued a minority position on compensation). The subjects did not move one cent in public. Yet, their judgments shifted dramatically on

post-experiment questionnaires both on the case they had discussed and also on new cases with new facts. It was a demonstration of strong resistance against any public agreement with minority views coupled with change at a latent level. That study also made us revise our compensation to include 'combat pay' for the confederate.

Mugny and his colleagues have elaborated the reasons for this reluctance to publicly move to the minority view by invoking considerations of social identity and the desire to belong (Butera and Mugny, 1995, 2001; Mugny et al., 1995). Other considerations have been ingroup and outgroup status. Almost all researchers demonstrate that there is more discomfort and anger (Philips, 2003) when the holder of a minority view is a member of one's ingroup rather than a member of an outgroup. And this discomfort may be one of [p. 366 ↓] the reasons why many studies show influence is greater from an ingroup member (Crano, 2000; David and Turner, 1996; Volpato et al., 1990). Discomfort may be a catalyst for further assessment of the message.

Such findings make it clear that it is not easy to maintain a minority viewpoint. Anger and the perceptions of being unintelligent (or worse) are a predictable consequence (Levine, 1989). However, people who persist in a minority view can also be accorded the perceptions of confidence and courage which can be assets in their attempts to influence (Nemeth and Chiles, 1988; Nemeth et al., 1977). This is still an uncharted area of research—why would anyone maintain a dissenting viewpoint given the consequences—but fortunately, it is starting to be investigated (De Dreu et al., 2000; LePine and Van Dyne, 2001).

Moving to Influence as a Stimulant to Thought

Learning that influence may occur at a private or latent level, even if not observable in public, helped us to question the paradigms used to study attitude change. The field conceptualised influence to be attitude change on a Likert scale. If the target moved along that scale to the position of the source, that was influence. Yet, we had increasing evidence that attitudes could change on other dimensions or on other related issues. But convincing researchers to broaden the definition of influence was not as easy as it

seemed. Yet, we were convinced that influence was far more subtle and deeper than movement on a Likert scale. It was perhaps a recognition of this possibility plus the fact that dissent had its power at the latent level that was to prime me for the observations that would come to dominate my professional life, and ironically to bring me back to the reasons why I decided to study social psychology.

My own early work on minority influence – the first five years post-PhD – was consumed with understanding when, how, and why minority views can prevail. A number of us who had studied with Moscovici had become a cohesive and determined minority ourselves and had documented and elaborated the power of the minority to persuade – to ‘win’. And we were finding converts to that field of research, initially with a wave of second-generation researchers. However, in 1974, my focus started to shift, much of it owing to an abiding interest in jury deliberations. Jury decisions were often the ‘task of choice’ in my experimental studies but it started to become a research interest in and of itself while at the University of Virginia.

The Supreme Court had recently ruled that a requirement of non-unanimity in jury decisions did not violate the defendant's constitutional rights (*Apodaca, Cooper, and Madden v. Oregon*, 1972; *Johnson v. Louisiana*, 1970). Oregon permitted a 10–2 verdict and Messrs Apodaca, Cooper and Madden were convicted by 10–2 or 11–1 verdicts. In Louisiana, crimes subject to hard labour were allowed a 9–3 verdict and Mr Johnson was convicted by such a vote. All of these individuals thus appealed their convictions based on a violation of the Sixth and Fourteenth Amendments. Under ‘equal protection of the law’, for example, their verdicts would have had to be unanimous had they been tried in another state. The court, however, ruled that their rights had not been violated and essentially argued that the verdicts would have been the same had unanimity been required.

At the time, there was evidence (Kalven and Zeisel, 1966) from real juries that the position held by a majority on the first ballot had a high likelihood (around 90 percent) chance of being the final verdict. Davis and his colleagues (Davis, 1989; Davis et al., 1984) had documented the power of majority views in mock jury decisions; in particular, a two-thirds majority seemed to fit the outcomes quite well. There was little evidence that the verdict differed as a result of a requirement of unanimity. But the issue was larger than that.

In reading the court's decision which, by the way, was rendered by a majority of five justices outvoting a minority of four [p. 367 ↓] justices – it was clear that there were broader considerations, those dealing with the process (not just the outcome) of decision making, theories about majority and minority influence, and community confidence. For example, the majority justices had a theory that the majority would not outvote a minority until 'reasoned discussion had ceased to have persuasive effect or to serve any other purpose – when a minority, that is, continues to insist upon acquittal without having persuasive reasons in support of its position' (*Johnson v. Louisiana*, 1970: 1624). The image of 'obstacle' and rogue dissenter reared its head again. By contrast, the minority opinion of four justices had a different theory and worried that polite and academic conversation might occur once the requisite votes were needed and this was 'no substitute for the earnest and robust argument necessary to reach unanimity' (*Johnson v. Louisiana*, 1970: 1647–1648). We decided to study the issue.

We first did an experimental study varying unanimity versus two-thirds majority rule on a first-degree murder case. We studied not only outcome (verdict) but also the process of decision making. In that study, we did a full Bales analysis, coding every comment in terms of who spoke, to whom it was addressed and which of the 12 categories (e.g., 'agreeing', 'giving information', 'asking for opinions') it exemplified. We also collected data on the robustness of the deliberation, on whether the minority was outvoted when the requisite number was reached as well as perceptions of justice and agreement with the verdict, all of which were issues in the court cases.

We then studied the issue in connection with the Law School at the University of Virginia. Third-year law students tried various cases in an actual courtroom with an actual judge and with witnesses. We brought the jury. Each jury of 12 was divided into two groups of six, one under a requirement of unanimity, the other a two-thirds majority. Deliberations were videotaped and analysed as in the experimental study. This way, we hoped for both experimental control and generalisability and fortunately, the findings proved to be in parallel, thus strengthening the findings. We learned that unanimity did not statistically alter the verdict but it did change the process. Under unanimity, deliberations were more robust, more fact oriented and people believed that justice had been better administered (Nemeth, 1977, 1984). However, we learned a great deal more.

Looking in detail at 40-minute deliberations of 60 or so groups, we started to notice something – call it a hunch or an insight. The groups with argumentation and debate, those where dissent was voiced and maintained, seemed to use more information, consider more options and, in our subjective judgment, had higher quality deliberations. That insight led to our first experimental study (Nemeth, 1976; Nemeth and Wachtler, 1983) where findings showed that subjects exposed to a disagreeing minority detected solutions that otherwise would have gone undetected. The task, a hidden figures array, was searched more fully and subjects saw the figure when embedded. They were not guessing; they were actively searching the array and finding ‘truth’ where it existed. This was in contrast to exposure to a disagreeing majority. In the latter situation, people concentrated on the figures suggested by the majority; they followed them exactly but did not find novel solutions. Such a finding corroborated our ‘hunch’ about the value of minority views and was presented to the first joint meeting of the American and European Societies of Experimental Social Psychology in Paris 1976.

That presentation had considerable impact. The audience seriously considered the possibility that minority views could influence thought and not just movement in attitude – public or private. I even think it influenced Moscovici’s conversion theory (1980) which, while still focusing on attitude change, theorised about the cognitive reasons for majority influence at the public level and minority influence at the private level. The former, he theorised, was a comparison process. The latter was a conversion process where people [p. 368 ↓] actively processed the minority’s arguments and position.

That first experimental study, conducted in 1973/1974 at Virginia, set the direction for many of our subsequent studies. And the findings provided optimism for raising the quality of thought and decision making. These ‘rogues’ might educate and stimulate us; they might make us smarter and wiser. It also dovetailed with the insights from that first jury study. Perhaps mostly, the jury work taught me where my interests lie. I didn’t really care who won, whether it was the majority or the minority. The possibility that we could construct groups that were wiser and smarter than the sum of the individuals was compelling. And it was in stark contrast to the findings that groups are defective, that they make poor decisions (Janis, 1972). At least they did not always have to do so.

After the initial work in Virginia, we conducted more studies, mostly at Berkeley, continually refining our conception of how majorities and minorities stimulate thought.

We already had information that those exposed to a minority view detected solutions that otherwise would have gone undetected (Nemeth, 1976; Nemeth and Wachtler, 1983). In a problem solving setting, those exposed to a minority utilised multiple strategies for problem solving (and performed better) while those exposed to a majority utilised the majority strategy to the detriment of other strategies (Nemeth and Kwan, 1987). There was also evidence of more originality of thought in response to minorities and more conventionality of thought in response to majorities (Nemeth and Kwan, 1985).

This early work convinced us that majorities and minorities stimulated different kinds of thinking and led to a theoretical formulation published in *Psychological Review* (Nemeth, 1986). Briefly, we hypothesised that majorities stimulate convergent thinking from the perspective posed by the majority. The thinking goes something like this: People exposed to a majority with a differing view are under stress and thus narrow the range of considerations. Further they assume the majority is correct and are motivated to assume that. As such, they focus on the issue or problem from the perspective of that majority in an attempt to understand why they take the position they do (and to find a reason to move to that position). By contrast, people exposed to a minority with a differing view assume the minority is in error. However, with consistency on the part of the minority, people come to reassess the situation and look at the issue anew. They don't assume the minority is correct but they are motivated to consider the issue more carefully since there must be a reason why the minority takes the position it does and, further, is sufficiently confident to maintain it.

Subsequent studies confirmed this set of hypotheses. We had evidence that minorities stimulated a search for information on all sides of the issue while majorities stimulated a search for information that corroborated the majority view (Nemeth and Rogers, 1996). There was better recall of information across categories in response to minority views (Nemeth et al., 1990). We even tested the strength of the theory by predicting when *majorities* might produce better performance. Remember the prediction is that majorities stimulate convergent focused thinking from their perspective.

In one study, we hypothesised that majorities could induce better performance IF it was a task where convergent thinking was useful and the perspective of the majority was appropriate. We chose the Stroop test, one of the few tasks where convergent thinking

is useful. People are shown a number of colour words which are printed in an ink of a different colour (e.g., the word red printed in green ink) and are asked to read the colour of *ink* as quickly and accurately as possible. This is a classic test of interference and is quite difficult as one often says the colour word (e.g., red) rather than the colour of ink (e.g., green). Here, if you can convergently focus on ink and not the name, this is adaptive. Conversely, if you convergently focus on name and not the ink, that is particularly maladaptive.

[p. 369 ↓]

The theory predicts that majorities stimulate convergent thinking from their perspective. The findings showed that when the majority focused on colour of *ink*, individuals were able to perform better on the Stroop test while a majority focus on *name* colour led to reduced performance. When it was a minority, their focus did not matter and performance was in between the two majority conditions (Nemeth et al., 1992). This study was essentially replicated with further evidence that minorities improve flexibility of performance (Peterson and Nemeth, 1996). Such findings gave us additional confidence about the theory since we could shift performance by altering the perspective of the majority and the nature of the task.

Research Strategy

The decision to do very simple problem-solving tasks rather than study the value of dissent in interacting groups was a deliberate one. While the origin of the idea was jury decision making and while our intended application was the role and value of dissent in small groups, we were also aware of the fact that interacting groups are very complex. They are difficult to study and it is difficult to establish cause and effect. Moreover, we wanted to make the point that the thinking, the performance and the decisions were not just different; they *were better*; they were correct. Attitudes are difficult to characterise as better, but if you show that people find correct solutions that otherwise would have gone detected, if you show they take in more information, consider more options and use all available strategies with resulting better performance and more originality, then it becomes easier to argue that groups profit from dissent, from minority views, especially since these effects occur even when those dissenting views are incorrect.

To some extent, this research strategy was also consistent with our theoretical perspective. Serge and I often spoke of the importance of making a clear point and being consistent, not just for personal integrity but also for influence. This was a bit contrary to graduate training where we complicated ideas, added variables and studied contingencies. However, in reflecting back on my early work on this issue, I had both a preference for simple (hopefully elegant) research designs and fundamentally believed that it is clarity, consistency and even simplicity that stimulates thought. The hope was that others would be stimulated to extend the thinking, to correct it, to elaborate on it and to show its boundary conditions, but the guiding theme of the value of dissent for divergent thought and clarity of position would remain. And we would welcome debate, for I was convinced that thought is stimulated by interaction, by discussion and, yes, even argument.

Impact and Application

Minority Influence and Social Psychology

A good deal of research developed the nature of cognitive activity in the realm of attitude change. Moscovici's (1980) conversion theory, for example, hypothesised quite different cognitive processes in response to a majority versus a minority source. The former created a comparison process where people identified with the majority and tried to 'fit in with their opinions or judgments'. Thus they often adopted the majority position – at least publicly – without scrutiny of the message. The latter created a conversion process – assuming the minority was consistent and confident – whereby people scrutinised the message. They wanted 'to see what the minority saw, to understand what it understood' (Moscovici, 1980: 215). This change, when it occurred, was deeper and longer lasting.

Competing theories arose (Mackie, 1987) which recognised that majorities also induced cognitive activity though, again, it was addressed to processing of the message and [p. 370 ↓] attitude change. A natural progression was for them to integrate the majority/minority source issue to the well-established research on peripheral versus central

(Petty and Caccioppo, 1981) or systematic versus heuristic (Chaiken, 1980) information processing and attitude change. As a result, we now know much more about such processing and resulting attitude change (see, generally, Hewstone and Martin, 2008; Wood et al., 1994).

What has occasionally concerned me (Nemeth, 2003), however, is that my own work has sometimes been misunderstood and assumed to predict attitude change – and thus included in the dialogue on message processing. Often, cognitive activity is construed as analysing the message of the source while my theory and work deals not so much with processing information as with thinking. Importantly, majorities stimulate people to think about the issue from their perspective. People adopt their framework, utilise their strategies and convince themselves of the truth of that position and way of thinking. Minorities do not just induce thought about their message; they induce thought about the issue. And importantly, they induce thought that is divergent, that considers multiple options only one of which is that suggested by the minority. People open their minds to information, to options, to creative possibilities.

This central idea was developed by a number of creative Italian researchers (Volpato et al., 1990) who demonstrated more original proposals as a result of exposure to minority views and extended the formulation by introducing the ingroup/outgroup nature of this influence. This is a direction further developed by Mugny and his colleagues in their cognitive elaboration model (Mugny and Papastamou, 1980; Mugny and Perez, 1991; Mugny et al., 1995) and, more recently by DeDreu and DeVries (1997) and Philips (2003).

When researchers have tried to integrate this theory into attitude change, it is a bit difficult because our hypothesis regarding divergent thinking doesn't predict what attitude will be adopted. It does suggest, however, that whatever attitude that is, will be better conceived, better understood and probably closer to reality. If you use an intervening variable such as the content and/or direction of thought, you can probably predict attitude change. But long ago, back at the beginning, 'winning' was not the main focus.

Apart from attitude change, there have also been some interesting recent developments related to the minority person's own cognitive activity. For example, Levine and

Russo (1995) show more divergent thought in preparation for being in a minority position. Others have connected the work to power. For example, there is evidence that those in power focus on a single target while those who are relatively powerless consider multiple sources of information (see Guinote, 2008). These studies have some interesting implications for the cognitive activity of the source and are consistent with hypotheses about differential cognitive activity associated with majority/minority status.

Somewhat ironically, those who have perhaps best understood and utilised our model have been in applied areas. Van Dyne and Saavedra (1996), for example, studied the role of dissent in work groups, finding that groups had improved decision quality when exposed to a minority perspective. They have recently replicated that finding and broadened the work to include value orientations of individualism/collectivism (Goncalo and Staw, 2006; Ng and Van Dyne, 2001). In a study of seven 'Fortune 500' top management teams, Peterson et al. (1998) found evidence that the most successful teams encouraged dissent in private meetings. And, in a study of US hospitals, Dooley and Fryxell (1999) found dissent related to high quality decision making teams.

Perhaps the greatest impact is evident in the applied areas of law and organisational culture. Both areas have a fundamental interest in the quality of decision making and creativity and these practical concerns have spurred an interest in possible mechanisms for achieving that. Dissent has been captured as one such vehicle (De Dreu and De Vries, [p. 371 ↓] 1997; Devine et al., 2001; Ford et al., 2008; Van Swol and Seinfeld, 2006).

Juries and Justice

In a study of the quality of the Supreme Court's own decisions, Gruenfeld (1995) found that the court's decisions were more integratively complex when there was dissent. In other words, there were more distinctions and integration of varying considerations when there was dissent. Another direct application is in the dialogue on procedural rules protecting dissent such as the requirement of unanimity in juries (Devine et al., 2001). And we were pleased to find the work cited in court cases such as an *amicus curiae* brief in support of considering race and gender in university admissions (*Grutter v. Bollinger, James et al.*) where the authors argue for the value of heterogeneity of views.

Some of this application, we believe, has come less from published articles than direct persuasion of lawyers and judges – a deliberate strategy on our part.

In the mid 1970s, partly as a consequence of the experimental work on the unanimity issue, there arose an opportunity to study psychology and law for a year at the Battelle Seattle Research Center, courtesy of Gordon Bermant. It was a time to actually learn about the law. We (Bermant et al., 1976) organised a conference and, more importantly, one participant – a judge from Portland, Oregon – agreed to host me for two weeks to study the jury system in Portland. Access to judges, prosecutors, public defenders, almost all files, ‘behind closed doors’ negotiations and dinner parties that were conducted like seminars, provided invaluable information about how things ‘worked’ and what issues were legally relevant. It also led to an invited address before the Oregon Bar Association in 1976. The topic for the 1,100 lawyers was, in part, the importance of unanimity for protection of dissenting views.

Some years later (2003), a similar talk before a group of researchers and judges from Australia and New Zealand led to discussion about the value of dissent as, at that time, both countries were considering ‘reform’ from unanimity to some form of majority rule in juries. While the promoters of this reform used the words ‘rogue’ and ‘obstacle’ to describe the dissenter, our presentation moved the discussion from efficiency to truth, justice and the possible value of dissent.

Corporate Cultures and Voice

The application of the dissent research to organisations has a similar trajectory but a broader one. There is now considerable evidence that the model of minority influence (Nemeth, 1986) is robust and well replicated in field settings (Peterson et al., 1998; Van Dyne and Saavendra, 1996). Further, a number of studies demonstrate that dissent increases creativity and better performance, at least under certain circumstances (De Dreu and De Vries, 1997; De Dreu et al., 2000; Ng and Van Dyne, 2001; Van Swol and Seinfeld, 2006). There are also three specific areas where the work has found application in organisations: the devils advocate, brain-storming, and corporate cultures.

The Devil's Advocate

I had always found it of interest that even when people believed that dissent had value, they still were trying to quash it or to find a mechanism that could have it 'both ways': keep its beneficial properties and yet be more palatable. Many observers, practitioners and researchers have struggled with the implications of dissent, primarily because they fear the frustration, lowered morale and 'slowing down' of the process. They thus often favour mechanisms such as devil's advocate, hoping they can avoid some of the 'downsides' of dissent.

Some research has found the technique to be of value though a number of studies compare it to having no alternatives presented [p. 372 ↓] (Katzenstein, 1996). Janis (1972) himself suggested devil's advocate as an antidote to 'groupthink' on the assumption that it would question the prevailing mode of thought and bring diverse viewpoints to the discussion. After years of working with dissent and thought stimulation, however, we never believed that you could clone the effects that easily. Serious reappraisal of a belief, we thought, required challenge and would be unlikely to occur in a role-playing setting. It was at least worth studying.

In one study (Nemeth et al., 2001b), we compared authentic dissent with devil's advocate in a mock jury decision-making experiment. Groups of four deliberated a personal injury case and decided on appropriate compensation by means of a series of votes and arguments. The position and arguments of the 'minority' remained the same. The only difference was whether or not the person was asked to play the 'devil's advocate' prior to the deliberation. Results showed that 'authentic dissent' (when no such instruction was given) led to more divergent thinking. The subjects generated novel thoughts that were on both sides of the issue. When that person was asked to play a 'devil's advocate', there not only was less thinking on both sides of the issue. There was evidence of 'cognitive bolstering'. Individuals generated thoughts that confirmed their initial position. Thus, we were able to show that not only was devil's advocate not as effective as authentic dissent but, further, it solidified the initial position.

In a second study (Nemeth et al., 2001a), we compared variations of devil's advocate. In the first study, the true position of the 'devil's advocate' was unknown. In the

second study, we simply varied whether it was unknown or known and, if known, if it was consistent with the position she advocated or inconsistent with the position she advocated (namely she agreed with the majority). A fourth condition was 'authentic dissent', namely no request to act as a devil's advocate. Surprisingly, the variations of devil's advocate did not matter; they did not differ significantly from one another. However, as predicted, none achieved the stimulating effect of authentic dissent. In this study, that stimulation took the form of creative solutions to the problem under discussion. Those exposed to authentic dissent generated more creative solutions than did those in any of the devil's advocate conditions. These findings have been replicated by Schultz-Hardt et al. (2002) demonstrating the value of majority/minority viewpoints in the context of information-seeking bias.

This issue of devil's advocate remains controversial as many companies and many researchers still argue for its value. Some well known CEOs have raised doubts about devil's advocate, not because it is ineffective but, rather, for the conflict that it creates. Dave Kelley of IDEO, perhaps the best-known design company in the world and known for its creative culture, argues that devil's advocates are 'naysayers' who can smother a fragile idea. He thus prefers using various nonconfrontational role-playing techniques (Kelley and Littman, 2005). Others, such as Harvard Business School professor, Dorothy Leonard, take our perspective, arguing for authentic dissent (Leonard and Swap, 1999). What becomes important is that there is a dialogue and debate, one we believe is served by clear alternatives and dissenting viewpoints.

Brainstorming

A technique long believed to enhance group creativity is that of brainstorming. This technique (Osborn, 1957), has persisted in practice even though the research shows little evidence of effectiveness. More precisely, brainstorming instructions increase the number of ideas in a group but it is usually less than the total number of ideas generated by the same number of individuals brainstorming alone (Brown and Paulus, 1996).

The instructions for brainstorming are fairly precise:

- 1. Quantity: come up with as many ideas as you can.
- 2. Do not criticise others' ideas.
- 3. Build on others' ideas.
- 4. Freewheeling is welcome.

The fact that brainstorming instructions do not achieve the level of a 'nominal' group (the sum of ideas made by the same number of individuals brainstorming alone) has led many researchers to focus on the 'losses' generated by groups. The culprits are usually motivational or coordination problems. In particular, Diehl and Stroebe (1987) review the available literature and conclude that one of the biggest problems is production blocking. People can't speak at the same time and, thus, ideas are often lost or not stated as a result. A good deal of work has focused on how to counter these losses – for example, having individuals write down their ideas and then discuss them as a group, using electronic brainstorming and so on. (Brown and Paulus, 1996).

Most of the work assumes that groups are less than the sum of their individuals and the aim is to counter those losses. The work on dissent and cognitive stimulation, however, suggests that groups can be better than the sum of the individuals (Nemeth and Nemeth-Brown, 2003), an orientation compatible with the extensive work done by Paulus and his colleagues (Dugosh et al., 2000; Paulus and Dzindolet, 1993; Paulus and Nijstad, 2003) who recognise that there can be a group synergy, that individuals can stimulate ideas in another. This has spawned research on the usage of computer interaction, facilitators, diversity and group goals, all of which figure in this potential synergy. The one thing that is missing is debate.

Note that, among the four 'rules' of brainstorming, there is the admonition not to criticise each others' ideas. The idea, of course, is that criticism will cause evaluation apprehension: people will be reluctant to express creative ideas; they won't 'free wheel' for fear of evaluation and risk. The dissent model, by contrast, predicts not only cognitive activity but also originality as a result of exposure to opposing views. This is a bit subtle in the brainstorming context. However, the role of debate and conflict might still be productive.

In a test of this possibility, we conducted a study in both the US and in France (Nemeth et al., 2004). The studies were identical. Individuals were given four instructions. In

one condition, it was the four rules as posited by Osborne. In the second condition, there was one exception. The admonition ‘not to criticise’ was replaced by ‘feel free to debate, even criticise’. Most researchers would have predicted that the ‘do not criticise’ instruction would lead to more ideas than a control group and that the ‘debate/criticise’ instruction would lead to fewer ideas than a control. In fact, we found the debate instruction led to significantly more ideas than a control, and it was even superior to the ‘do not criticise’ (though not significantly so). These findings were reflected in both the US and France.

Corporate Cultures

Given the interests of business schools on the topic of innovation, the work on dissent was a natural fit. While some learned of the formulation (Nemeth, 1986) in a psychological journal, it was an article directly pertaining to corporate cultures that had the greatest impact (Nemeth, 1977). This article was a direct consequence of an invited address at a conference on Knowledge and the Firm at the Haas School of Business and, much like the legal research, the article required a year to learn about company practices. We came to the conclusion that most organisations try to create cohesion, harmony and alignment with a company vision. They want creativity and innovation but do not embrace the idea of welcoming dissent. In fact, they reward loyalty rather than innovation. And there are many business gurus who are happy to help them achieve the benefits of a ‘cultlike’ culture (Collins and Porras, 1994).

Given the years of documenting the value of differing views, the potential ‘downsides’ of morale and cohesion were evident. After all, we and others repeatedly found that [p. 374 ↓] majority views stimulate convergent thinking from their perspective. This is useful provided that the majority perspective was the correct or best one. Thus company practices such as recruiting those who ‘fit’ the organisation, socialising and interaction, and ejecting dissent ‘as a virus’ – all of which are argued to characterise ‘visionary’ companies (Collins and Porras, 1994) – seemed relevant for identification, for a sense of belonging and morale but not necessarily for performance and certainly not for innovation. Cults use similar practices (O'Reilly and Chatman, 1996).

From our perspective, the problem seemed to be that such convergence of thought and action is useful to the extent that they are on the *right* page, not just the *same* page. Remember the Stroop study (Nemeth et al., 1992)? The kinds of thoughts stimulated by such practices, we argued, are powerful for implementation of an idea and depend on the value of that idea but they are not conducive to developing or changing an idea, to recognising a changing market or new opportunities. Rather, it is dissent or at least the open airing of competing views that could do this (Nemeth, 1997).

This article, peppered with examples of corporate cultures, was a counter to much of the work arguing for a cult-like culture which, ironically, was often seen as compatible with innovation. As such, this work on the value of dissent (Nemeth, 1997) has become part of the dialogue on the role of dissent, debate, 'voice' and conflict in organisations (Amason, 1996; Ford et al., 2008; Van Dyne et al., 2003). While it would be comforting to think that people are learning to 'welcome' dissent, I suspect that it is more often a stimulant for discussion than taken at face value. The message is still difficult for managers who often want to control when and where innovation occurs. And the recurrent themes of chaos, wasted time and reduced morale still remain.

The issues involved in corporate cultures are complex and varied, ones I knew would require further education on my part, somewhat similar to that required for the work on the law. That led to a decision to spend another year immersing myself in the issues. I took an unpaid leave of absence in 2005/2006 to teach at the London Business School – another humbling experience of learning. But being a part of the voices for the value of diversity and dissent in the workplace (Detert and Edmondson, 2008; Morrison et al., 2003; Van Dyne et al., 2002) is an opportunity for impact on social issues. It is also an opportunity for complicating my own research, a result whenever one tries to study applied issues.

I learned that it is difficult to balance performance, profitability and innovation. It is difficult to manage creative individuals and the creative process. But it is also difficult to persuade people to speak up and to manage an 'unwanted truth' (Edmondson and Munchus, 2007; Morrison and Milliken, 2000). The problems remain: people fear that speaking up will not serve any purpose save their own branding as a 'nuisance'. No one finds it easy to 'welcome' disagreement.

Conclusion

Over the course of these decades studying influence in the laboratory and, more recently, in naturalistic settings, I am convinced of a few things. People are loath to change their minds easily, the attitude change literature notwithstanding. Serge was right. It requires exposure to a differing view, not just in content but in contrast, discussion and debate.

I am also happily convinced that authenticity is important. Most of our work documents the value of authentic dissent. Dissenting for the sake of dissenting is not useful. It is also not useful if it is ‘pretend dissent’ – for example, if role-played. It is not useful if motivated by considerations other than searching for the truth or the best solutions. But when it is authentic, it stimulates thought, it clarifies and it emboldens.

Finally, it is clear to me that it is still difficult to convince people of the value of diversity and dissent. They accept the principle on the [p. 375 ↓] surface – it sounds nice, democratic and tolerant – but, in fact, people get quickly irritated by a dissenting view that persists and fear the lowered morale, the lack of ‘harmony’ and a loss of control by ‘welcoming’ dissent. Thus, we continually find attempts to denigrate it or to contain it. People are encouraged to ‘role play’ their ideas instead of stating them clearly; they are asked to ‘fit in’, to be on the same page, to not make waves and to be in line with the leader’s views or the company vision. They are made to fear repercussions, including being marginalised by gossip or ridicule. I often think that it is our differences that make us interesting as human beings and it is in our differences and our willingness to embrace them that we learn and grow, that at least we *think*.

What I hope is that we start to recognise the courage of minority voices and the value of the open airing of competing views (John Stuart Mill, 1859) and that we achieve some clear understanding of the role of trust that allows the passionate interchange to occur. What I also hope is that we put less emphasis on ‘winning’, persuading and manipulating others and return to the ways in which interaction clarifies, educates and elevates us.

Acknowledgements

This manuscript was supported by the Kauffman Foundation and by the Institute for Industrial Relations at UC Berkeley, support which is gratefully acknowledged. Requests for reprints should be addressed to Professor Charlan Nemeth, Department of Psychology, University of California, Berkeley, CA 94720–1650.

I owe a debt of gratitude to the wonderful graduate students with whom I collaborated at Chicago, Virginia, British Columbia and Berkeley. Special ones who worked with me for years include Jeff Endicott and Joel Wachtler (at Chicago and Virginia), Jack Goncalo, John Rogers, and Keith Brown (at Berkeley). They taught me as much as I taught them. Wonderful other graduate students include: Cindy Chiles, Matt Feinberg, Linda George, Julie Kwan, Ofra Mayseless, Kathy Mosier, Margaret Ormiston, Rhonda Pajak, Randall Peterson, Jeff Sherman and Elaine Wong.

Charlan JeaneNemeth

References

Amason, A.C. Distinguishing the effects of functional and dysfunctional conflict on strategic decision making: Resolving a paradox for top management teams . Academy of Management Journal , 39, 123–148. (1996) <http://dx.doi.org/10.2307/256633>

Apodaca, Cooper, and Madder v. Oregon, 406, U.S., 404 (1972).

Asch, S.E. Studies of independence and conformity: A minority of one against a unanimous majority . Psychological Monographs , 70, Whole no. 416. (1956)

Berger, J.M., Fisek, M.H., Norman, R.Z. and Zelditch, M. (1977) Status Characteristics in Social Interactions: An Expectation States Approach . New York: Elsevier.

Berman, G., ed. , Nemeth, C., ed. and Vidmar, N. (eds) (1976) Psychology and the Law: Research Frontiers . Lexington: D.C. Heath and Co.

Brown, V. and Paulus, P.B. A simple dynamic model of social factors in group brainstorming . *Small Group Research* , 27, 91–114. (1996) <http://dx.doi.org/10.1177/1046496496271005>

Butera, F. and Mugny, G. Conflict between incompetences and influence of a low-expertise source in hypothesis testing . *European Journal of Social Psychology* , 25, 457–462. (1995) <http://dx.doi.org/10.1002/ejsp.2420250408>

Butera, F., ed. and Mugny, G. (eds) (2001) *Social Influence in Social Reality: Promoting Individual and Social Change* . Ashland: Hogrefe and Huber Publishers.

Chaiken, S. Heuristic versus systematic information processing and the use of source versus message cues in persuasion . *Journal of Personality and Social Psychology* , 39, 752–766. (1980) <http://dx.doi.org/10.1037/0022-3514.39.5.752>

Collins, J.C. and Porras, J.I. (1994) *Built to Last: Successful Habits of Visionary Companies* . New York: HarperCollins.

Crano, W.D. Social influence: effects of leniency on majority and minority-induced focal and indirect attitude change . *Revue Internationale de Psychologie Sociale* , 15, 89–121. (2000)

David, B. and Turner, J.C. Studies in self-categorization and minority conversion: Is being a member of the out-group an advantage? *British Journal of Social Psychology. Special Issue: Minority Influences* , 35, 179–199. (1996)

David, B. and Turner, J.C. (2001) Self-categorization principles underlying majority and minority influence . In J.P. Forgas, ed. and K.D. Williams (eds), *Social Influence: Direct and Indirect Processes* . Philadelphia: Psychology Press.

Davis, J.H. Psychology and law: The last 15 years . *Journal of Applied Social Psychology* , 19, 199–230. (1989) <http://dx.doi.org/10.1111/j.1559-1816.1989.tb00052.x>

Davis, J.H., Tindale, R.S., Nagao, D.H., Hinsz, V.B. and Robertson, B. Order effects in multiple decisions by groups: A demonstration with mock juries and trial procedures . *Journal of Personality and Social Psychology* , 44, 20–33. (1984)

De Dreu, C.K.W. and De Vries, N.K. (1997) Minority dissent in organizations . In C.K.W. De Dreu, ed. and E. Van de Vliert (eds), Using Conflict in Organizations , pp. 72–86. Thousand Oaks: Sage. <http://dx.doi.org/10.4135/9781446217016>

De Dreu, C.K.W., De Vries, N.K., Franssen, H. and Altink, W.M.M. Minority dissent in organizations: Factors influencing willingness to dissent . Journal of Applied Social Psychology , 30, 2451–2466. (2000) <http://dx.doi.org/10.1111/j.1559-1816.2000.tb02445.x>

Detert, J.R. and Edmondson, A.C. (2008) Silent Saboteurs: How Implicit Theories of Voice Inhibit the Upward Flow of Knowledge in Organizations . Working Knowledge , Harvard Business School, December 2008.

Devine, D.J., Clayton, L.D., Dunford, B.B., Seying, R. and Pryce, J. Jury decision making: 45 years of empirical research on deliberating groups . Psychology, Public Policy, and Law , 7, 622–727. (2001) <http://dx.doi.org/10.1037/1076-8971.7.3.622>

Diehl, M. and Stroebe, W. Productivity loss in idea-generating groups: Tracking down the blocking effect: Toward the solution of a riddle . Journal of Personality and Social Psychology , 53, 497–509. (1987) <http://dx.doi.org/10.1037/0022-3514.53.3.497>

Dooley, R.S. and Fryxell, G.E. Attaining decision quality and commitment from dissent: The moderating effects of loyalty and competence in strategic decision-making teams . Academy of Management Journal , 42, 389–402. (1999) <http://dx.doi.org/10.2307/257010>

Dugosh, K.L., Paulus, P.B., Roland, E.J. and Yang, H. Cognitive stimulation in brainstorming . Journal of Personality and Social Psychology , 79, 722–735. (2000) <http://dx.doi.org/10.1037/0022-3514.79.5.722>

Edmondson, V.C. and Munchus, G. Managing the unwanted truth: A framework for dissent strategy . Journal of Organizational Change Management , 20, 747–760. (2007) <http://dx.doi.org/10.1108/09534810710830993>

Faucheu, C. and Moscovici, S. The style of behavior of a minority and its influence on majority responses . Bulletin Du C.E.R.P . 16, 337–361. (1967)

Ford, J.D., Ford, L.W. and D'Amelio, A. Resistance to change: The rest of the story . Academy of Management Review , 33, 362–377. (2008) <http://dx.doi.org/10.5465/AMR.2008.31193235>

Forgas, J.P., ed. and Williams, K.D. (eds) (2001) Social Influence: Direct and Indirect Processes: The Sydney Symposium of Social Psychology . New York: Psychology Press.

Giles, H., Bourhis, R.Y. and Davies, A. (1979) Prestigious speech styles: The imposed norm and inherent value hypothesis . In W. McCormack, ed. and S. Wurm (eds), Language in Anthropology IV: Language in Many Ways , pp. 307–348. The Hague: Mouton.

Goncalo, J.A. and Staw, B.M. Individualism-collectivism and group creativity . Organizational Behavior and Human Decision Processes , 100, 96–109. (2006) <http://dx.doi.org/10.1016/j.obhdp.2005.11.003>

Gruenfeld, D.H. Status, ideology, and integrative complexity on the U.S. Supreme Court: Rethinking the politics of political decision making . Personality and Social Psychology Bulletin , 68, 1013–1022. (1995)

Grutter v. Bollinger, James et al., 539 U.S. 306 (2003).

Guinote, A. Power and affordances: When the situation has more power over powerful than powerless individuals . Journal of Personality and Social Psychology , 95, 237–252. (2008) <http://dx.doi.org/10.1037/a0012518>

Hewstone, M. and Martin, R. (2008) Social influence . In M. Hewstone, ed. , W. Stroebe, ed. and K. Jonas (eds), Introduction to Social Psychology , 4th Edition, pp. 216–243. Malden: Blackwell Publishing.

Hovland, C.I., Janis, I.L. and Kelley, H.H. (1953) Communication and Persuasion: Psychological Studies of Opinion Change . New Haven: Yale University Press.

Janis, I.L. (1972) Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascos . Oxford: Houghton Mifflin.

Johnson v. Louisiana, 406 U.S. 356 (1972).

Kalven, H., Jr. and Zeisel, H. (1966) *The American Jury*. Boston: Little, Brown.

Katzenstein, G. The debate on structured debate: Toward a unified theory . *Organizational Behavior and Human Decision Processes* , 66, 316–332. (1996) <http://dx.doi.org/10.1006/obhd.1996.0059>

Kelley, T. and Littman, J. (2005) *The Ten Faces of Innovation: IDEO's Strategies for Beating the Devil's Advocate and Driving Creativity Throughout your Organisation* . New York: Currency, Doubleday.

Kiesler, C.A. and Kiesler, S.B. (1970) *Conformity* . New York: Addison Wesley Publishing Company.

Latané, B. The psychology of social impact . *American Psychologist* , 36, 343–356. (1981)

Leonard, D.A. and Swap, W.C. (1999) *When Sparks Fly: Igniting Creativity in Groups* . Boston: Harvard Business School Press.

LePine, J.A. and Van Dyne, L. Voice and cooperative behavior as contrasting forms of contextual performance: Evidence of differential relationships with big five personality characteristics and cognitive ability . *Journal of Applied Psychology* , 86, 325–336. (2001) <http://dx.doi.org/10.1037/0021-9010.86.2.326>

Levine, J.M. (1989) Reaction to opinion deviance in small groups . In P.B. Paulus (ed.), *Psychology of Group Influence* , 2nd Edition, pp. 187–231. Hillsdale: Lawrence Erlbaum Associates.

Levine, J.M. and Russo, E. Impact of anticipated interaction on information acquisition . *Social Cognition* , 13, 293–317. (1995) <http://dx.doi.org/10.1521/soco.1995.13.3.293>

Maass, A., West, S.G. and Cialdini, R.B. (1987) Minority influence and conversion . In C. Hendrick (ed.), *Group Processes* , pp. 55–79. Thousand Oaks: Sage.

Mackie, D.M. Systematic and nonsystematic processing of majority and minority persuasive communications . *Journal of Personality and Social Psychology* , 53, 41–52. (1987) <http://dx.doi.org/10.1037/0022-3514.53.1.41>

Mill, J.S. (1859) *On Liberty* . New York: Penguin.

Morrison, E.W. and Milliken, F.J. Organizational silence: A barrier to change and development in a pluralistic world . *Academy of Management Review* , 25, 706–725. (2000)

Morrison, E.W., Milliken, F.J., Dyne, L., Ang, S., Botero, I.C. and Bowen, F., Speaking up, remaining silent: The dynamics of voice and silence in organizations . *Journal of Management Studies* , 40, 1353–1568. (2003) <http://dx.doi.org/10.1111/1467-6486.00383>

Moscovici, S. (1980) Toward a theory of conversion behavior . In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, Vol. 13 . New York: Academic Press.

Moscovici, S. and Faucheu, C. (1972) Social influence, conforming bias, and the study of active minorities . In L. Berkowitz (ed.), *Advances in Experimental Social Psychology* , 6, 149–202. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960027-1>

Moscovici, S. and Lage, E. Studies in social influence III: Majority versus minority influence in a group . *European Journal of Social Psychology* , 6, 149–174. (1976) <http://dx.doi.org/10.1002/ejsp.2420060202>

Moscovici, S., Lage, E. and Naffrechoux, M. Influence of a consistent minority on the responses of a majority in a color perception task . *Sociometry* , 32, 365–380. (1969) <http://dx.doi.org/10.2307/2786541>

Moscovici, S. and Nemeth, C. (1974) *Social Influence: II. Minority Influence* . Oxford: Rand McNally.

Mugny, G. (1982) *The Power of Minorities* . London: Academic Press.

Mugny, G., Butera, F., Sanchez-Mazas, M. and Pérez, J.A. (1995) Judgments in conflict: The conflict elaboration theory of social influence . In B. Boothe, ed. , R. Hirsig, ed. , A. Helminger, ed. , B. Meier, ed. and R. Volkart (eds), Perception-Evaluation-Interpretation , pp. 160–168. Ashland: Hogrefe and Huber Publishers.

Mugny, G. and Papastamou, S. When rigidity does not fail: Individualization and psychologization as resistances to the diffusion of minority innovations . European Journal of Social Psychology , 10, 43–61. (1980) <http://dx.doi.org/10.1002/ejsp.2420100104>

Mugny, G. and Perez, J.A. (1991) The Social Psychology of Minority Influence . New York: Cambridge University Press.

Nemeth, C. (1976) A comparison between conformity and minority influence . Paper presented to the International Congress of Psychology, Paris, France.

Nemeth, C. Interactions between jurors as a function of majority vs. unanimity decision rules . Journal of Applied Social Psychology , 7, 38–56. (1977). <http://dx.doi.org/10.1111/j.1559-1816.1977.tb02416.x>

Nemeth, C., Brown, K. and Rogers, J. Devil's advocate versus authentic dissent: Stimulating quantity and quality . European Journal of Social Psychology , 31, 707–720. (2001a) <http://dx.doi.org/10.1002/ejsp.58>

Nemeth, C. and Chiles, C. Modeling courage: The role of dissent in fostering independence . European Journal of Social Psychology , 18, 275–280. (1988) <http://dx.doi.org/10.1002/ejsp.2420180306>

Nemeth, C., Mayseless, O., Sherman, J. and Brown, Y. Exposure to dissent and recall of information . Journal of Personality and Social Psychology , 58, 429–437. (1990) <http://dx.doi.org/10.1037/0022-3514.58.3.429>

Nemeth, C., Mosier, K. and Chiles, C. When convergent thought improves performance: Majority versus minority influence . Personality and Social Psychology Bulletin , 18, 139–144. (1992) <http://dx.doi.org/10.1177/0146167292182004>

Nemeth, C. and Rogers, J. Dissent and the search for information . British Journal of Social Psychology. Special Issue: Minority Influences , 35, 67–76. (1996)

Nemeth, C., Swedlund, M. and Kanki, B. Patterning of a minority's responses and their influence on the majority . European Journal of Social Psychology , 4, 53–64. (1974) <http://dx.doi.org/10.1002/ejsp.2420040104>

Nemeth, C. and Wachtler, J. Creating the perceptions of consistency and confidence: A necessary condition for minority influence . Sociometry , 37, 529–540. (1974) <http://dx.doi.org/10.2307/2786425>

Nemeth, C., Wachtler, J. and Endicott, J. Increasing the size of the minority: Some gains and some losses . European Journal of Social Psychology , 7, 15–27. (1977) <http://dx.doi.org/10.1002/ejsp.2420070103>

Nemeth, C.J. (1984) Processus de groupe et jurys: Les Etats-Unis et la France . In S. Moscovici (ed.), Psychologie Sociale , pp. 229–251. Paris: Presses Universitaires de France.

Nemeth, C.J. Differential contributions of majority and minority influence . Psychological Review , 93, 23–32. (1986) <http://dx.doi.org/10.1037/0033-295X.93.1.23>

Nemeth, C.J. Managing innovation: when less is more . California Management Review , 40, 59–74. (1997)

Nemeth, C.J. Minority dissent and its 'hidden' benefits . New Review of Social Psychology , 2, 21–28. (2003)

Nemeth, C.J., Connell, J.B., Rogers, J.D. and Brown, K.S. Improving decision making by means of dissent . Journal of Applied Social Psychology , 31, 48–58. (2001b) <http://dx.doi.org/10.1111/j.1559-1816.2001.tb02481.x>

Nemeth, C.J. and Kwan, J.L. Originality of word associations as a function of majority vs minority influence . Social Psychology Quarterly , 48, 277–282. (1985) <http://dx.doi.org/10.2307/3033688>

Nemeth, C.J. and Kwan, J.L. Minority influence, divergent thinking and detection of correct solutions . *Journal of Applied Social Psychology* , 17, 788–799. (1987) <http://dx.doi.org/10.1111/j.1559-1816.1987.tb00339.x>

Nemeth, C.J. and Nemeth-Brown, B. (2003) Better than individuals? the potential benefits of dissent and diversity for group creativity . In P.B. Paulus, ed. and B.A. Nijstad (eds), *Group Creativity: Innovation Through Collaboration* , pp. 63–84. New York: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780195147308.003.0004>

Nemeth, C.J., Personnaz, B., Personnaz, M. and Goncalo, J.A. The liberating role of conflict in group creativity: A study in two countries . *European Journal of Social Psychology* , 34, 365–374. (2004) <http://dx.doi.org/10.1002/ejsp.210>

Nemeth, C.J. and Wachtler, J. Creative problem solving as a result of majority vs. minority influence . *European Journal of Social Psychology* , 13, 45–55. (1983) <http://dx.doi.org/10.1002/ejsp.2420130103>

Ng, K.Y. and Van Dyne, L. Individualism-collectivism as a boundary condition for effectiveness of minority influence in decision making . *Organizational Behavior and Human Decision Processes* , 84, 198–225. (2001) <http://dx.doi.org/10.1006/obhd.2000.2927>

O'Reilly, C.A. and Chatman, J.A. Culture as social control: Corporations, cults and commitment . *Research in Organizational Behavior* , 18, 157–200. (1996)

Osborn, A.F. (1957) *Applied Imagination* , 1st Edition. New York: Scribner.

Paulus, P.B. and Dzindolet, M.T. Social influence processes in group brainstorming . *Journal of Personality and Social Psychology* , 64, 575–586. (1993) <http://dx.doi.org/10.1037/0022-3514.64.4.575>

Paulus, P.B., ed. and Nijstad, B.A. (eds) (2003) *Group Creativity: Innovation Through Collaboration* . New York: Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780195147308.001.0001>

Peterson, R.S. and Nemeth, C.J. Focus versus flexibility: majority and minority influence can both improve performance . *Personality and Social Psychology Bulletin* , 22, 14–23. (1996) <http://dx.doi.org/10.1177/0146167296221002>

Peterson, R.S., Owens, P.D., Tetlock, P.E., Fan, E.T. and Martorana, P. Group dynamics in top management teams: Groupthink, vigilance, and alternative models of organizational failure and success . *Organizational Behavior and Human Decision Processes*. Special Issue: Theoretical Perspectives on Groupthink: A Twenty-Fifth Anniversary Appraisal , 73, 272–305. (1988) <http://dx.doi.org/10.1006/obhd.1998.2763>

Petty, R.E. and Cacioppo, J.T. (1981) Attitudes and Persuasion — Classic and Contemporary Approaches . Dubuque: William C. Brown.

Phillips, K. The effects of categorically based expectations on minority influence: the importance of congruence . *Personality and Social Psychology Bulletin* , 29, 3–12. (2003) <http://dx.doi.org/10.1177/0146167202238367>

Schachter, S. Deviation, rejection, and communication . *Journal of Abnormal and Social Psychology* , 46, 190–207. (1951) <http://dx.doi.org/10.1037/h0062326>

Schulz-Hardt, S., Jochims, M. and Frey, D. Productive conflict in group decision making: Genuine and contrived dissent as strategies to counteract biased information seeking . *Organizational Behavior and Human Decision Processes* , 88, 563–586. (2002) <http://dx.doi.org/10.1016/S0749-5978%2802%2900001-8>

Van Dyne, L., Ang, S. and Botero, I.C. Conceptualizing employee silence and employee voice as multidimensional constructs . *Journal of Management Studies*. Special Issue: Speaking Up, Remaining Silent: The Dynamics of Voice in Organizations , 40, 1359–1392. (2003)

Van Dyne, L., Jehn, K.A. and Cummings, A. Differential effects of strain on two forms of work performance: Individual employee sales and creativity . *Journal of Organizational Behavior* , 23, 57–74. (2002) <http://dx.doi.org/10.1002/job.127>

Van Dyne, L. and Saavedra, R. A naturalistic minority influence experiment: Effects of divergent thinking, conflict and originality in work-groups . British Journal of Social Psychology. Special Issue: Minority Influences , 35, 151–167. (1996)

Van Swol, L.M. and Seinfeld, E. Differences between minority, majority, and unanimous group members in the communication of information . Human Communication Research , 32, 178–197. (2006) <http://dx.doi.org/10.1111/j.1468-2958.2006.00008.x>

Volpato, C., Maass, A., Mucchi-Faina, A. and Vitti, E. Minority influence and social categorization . European Journal of Social Psychology , 20, 119–132. (1990) <http://dx.doi.org/10.1002/ejsp.2420200204>

Wood, W., Lundgren, S., Ouellette, J.A., Busceme, S. and Blackstone, T. Minority influence: A meta-analytic review of social influence processes . Psychological Bulletin , 115, 323–345. (1994) <http://dx.doi.org/10.1037/0033-2909.115.3.323>

<http://dx.doi.org/10.4135/9781446249222.n44>

Handbook of Theories of Social Psychology

Social Identity Theory

Contributors: Naomi Ellemers & S. Alexander Haslam

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Social Identity Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n45>

Print pages: 379-399

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n45>

[p. 379 ↓]

Chapter 45: Social Identity Theory

Abstract

Social identity theory is a “grand” theory. Its core premise is that in many social situations people think of themselves and others as *group members*, rather than as unique individuals. The theory argues that social identity underpins intergroup behavior and sees this as qualitatively distinct from interpersonal behavior. It delineates the circumstances under which social identities are likely to become important, so that they become the primary determinant of social perceptions and social behaviors. The theory also specifies different strategies people employ to cope with a devalued social identity. Social identity theory is a truly social psychological theory, in that it focuses on *social context* as the key determinant of self-definition and behavior. People's responses are thus understood in terms of subjective beliefs about different groups and the relations between them, rather than material interdependencies and instrumental concerns, objective individual and group characteristics, or individual difference variables. After its initial formulation as a “theory of intergroup conflict” in the 1970s, the theory has undergone many expansions, refinements, and updates. It has inspired a large body of research, and has been applied to inform the analysis of a range of issues and problems in group dynamics and intergroup relations.

Introduction

Since it was first proposed in the 1970s, social identity theory (SIT) has been considered one of the major theories in social psychology. It is consistently represented in textbooks and readers, has been used as a theoretical framework in countless empirical investigations, and has informed the analysis of a range of topics in group processes and intergroup relations (for an overview see Haslam et al., 2010a). If this

level of representation in the literature indicates the continued relevance of and steady interest in SIT, it also results from the continuous controversy and sometimes heated debate associated with this theoretical approach. Awareness that SIT has enthusiastic supporters as well as strong opponents merits a careful and critical consideration of the value of this theoretical approach. There should be no doubt, however, about the importance of being *informed* about SIT and the work it has inspired, if only to be able to understand ongoing scientific debates, and to allow the [p. 380 ↓] reader to place what represents a considerable part of the social psychological literature in an appropriate intellectual context. This chapter aims to provide such information, by describing the origins and further development of SIT and its main applications, as well as providing an overview of the main issues that have emerged as topics of debate over the years.

Origins of the Theory

As was the case with many other social psychologists of his generation, the scientific work of Henri Tajfel – a Jewish survivor of World War II of Polish birth – was inspired by his personal experiences of discrimination and intergroup conflict. In his early writings he explains that he was motivated to understand how people who had been living together as neighbors, colleagues, and friends could come to see each other as dangerous enemies even when there were no rational or objective reasons to do so. However, rather than take a field-study approach (e.g., Sherif, 1967), he sought to understand these issues by using scientific rigor to study groups in the laboratory and by exploring basic social cognitive processes which had been shown to be important in some of his earlier studies on object categorization (Tajfel, 1969).

This resulted in a series of experiments that later became known as the “minimal group studies” (Tajfel et al., 1971). Participants in these studies were informed that they had been assigned to one of two groups on the basis of an irrelevant criterion, or on the basis of chance. They did not know who else was present, they could not see or interact with others, and it was made clear that the choices they made could not affect their own outcomes in any way. Their task was then to allocate points to one member of their own group (not themselves), and one member of the other group. These “minimal” conditions were originally intended to form a baseline or control condition for further studies. As none of the known reasons to differentiate between a member of one's own

group and a member of another group were present, participants were expected to divide the points equally between them.

The historic significance of these studies lay in the observation that even these very minimal conditions proved sufficient to induce ingroup favoritism: the tendency to systematically allocate more points to a member of one's own group than to a member of another group. This effect later became known as the "mere categorization" effect – suggesting that the mere act of categorizing individuals into groups made people think of themselves and others in terms of "us" and "them," and was sufficient to induce them to behave differently towards ingroup and outgroup members.

Importantly, these findings were at odds with the scientific understanding of the time, which was informed largely by realistic conflict theory (RCT). RCT suggests that conflicts between members of different groups arise from competition over scarce resources (Sherif, 1967), but, as noted above, these were conspicuously absent from the minimal group paradigm. The provocative nature of the findings that emerged from the minimal group studies inspired a large body of research that attempted to examine alternative explanations for the mere categorization effect (Diehl, 1990; Rabbie et al., 1989). At the same time, these early findings informed the further development of SIT. After a series of publications in which Tajfel introduced the concept of social identity, and explained how the minimal group studies pointed to the fact that people sometimes behave as group members rather than as individuals (Tajfel, 1974, 1975, 1978a, 1978b, 1978c), he formulated SIT together with his Bristol colleague, John Turner, presenting it as "a theory of intergroup conflict" (Tajfel and Turner, 1979; see also Tajfel, 1982).

Basic Principles

The concept of social identity is defined as "that part of an individual's self-concept [p. 381 ↓] which derives from his knowledge of his membership of a social group (or groups) together with the emotional significance attached to that membership" (Tajfel, 1974: 69). The main aim of SIT is to understand and explain how people can come to adopt and behave in terms of such social (rather than personal) identities. When do people think of themselves in terms of "we" instead of "I"? Why is it important to know whether others can be seen as representing "us" or "them"? How does this impact upon

our feelings, thoughts, and behaviors? SIT tries to answer these questions, by pointing to the implications of social identity for the perceptions and behaviors of individuals, and examining the way in which this impacts on social relations between individuals and groups.

The basic principles of the theory address three main issues. First, they describe the *psychological processes* that explain how people's social identities are different from their personal identities. Second, they distinguish between different *strategies* people can use to derive a positive social identity. And third, they specify the key *characteristics of the social structure* that determine which of these strategies is most likely to be used in any given case.

Psychological Processes

Social categorization is the process through which separate individuals are clustered into groups. Social categorization is seen as a common and functional psychological process that provides a way of responding to complex social situations. Thinking of individuals in terms of a limited number of social categories provides a way of organizing socially relevant information, and helps in the process of both understanding and predicting behavior. When individuals are categorized into the same group, they are thought to share some central group-defining feature, which distinguishes them from others who do not possess this feature (Tajfel, 1978a). For instance, just as different pieces of furniture in a room can be classified as tables or chairs, so people in a school can be classified as students or teachers. As a result of such classifications, we tend to focus on similarities between individuals within the same category, and see them as interchangeable elements that share some representative common characteristics (e.g., a specific profession, religion, or national citizenship). At the same time, we accentuate differences between individuals who are classified into different categories (e.g., psychologists or economists), as a way of clarifying the meaning of the situation (Tajfel and Wilkes, 1963; Tajfel, 1978a, 1978b, 1978c). Thus, when people are categorized into groups, they come to be seen in terms of characteristic group features that define their social identities (economists use mathematical models), while neglecting individual traits which define their uniqueness (this particular economist interviews people about their emotions).

Social comparison is the process through which characteristic group features are interpreted and valued. Because there is no objective standard that enables us to assess the worth of different groups, we tend to decide whether a group is “good” or “bad” at something, by comparing the characteristics (e.g., traits, attitudes, behaviors) that are seen to define them to the characteristics ascribed to other groups (Tajfel, 1978b, 1978c). Thus, in parallel to interpersonal comparisons that may help determine individual worth (Festinger, 1954), groups and their features can also be evaluated by comparing them with other groups and their defining features (see also Levine and Moreland, 1987). For instance, sociologists may see themselves as relatively more “scientific” than historians, but as less “scientific” than physicists. The constellation of different group traits and how these compare to the traits of most other groups in that context determines the social status or perceived prestige of that group. Where social categorization determines how individuals are classified into groups, social comparisons define the ways in which each group is distinguished from relevant other groups.

[p. 382 ↓]

Social identification speaks to one key reason why groups of people are different from object categories: the fact that the *self* can also be seen as belonging to a social group (Tajfel, 1974, 1978a, 1978b). That is, just as pieces of furniture can be categorized as chairs or tables, so individuals can be socially categorized as men or women. And just as the comparison between chairs and tables leads to the conclusion that tables are generally larger than chairs, so the comparison between men and women leads to the conclusion that men are generally taller than women. But what makes social categories different from object categories is the process of social identification: the realization that the *self* is included in some social categories, and excluded from others. It is impossible, for instance, to categorize and compare men and women without realizing that one of these categories includes the *self*. Thus, when specific features are associated with a social group, or when these features are valued in a certain way, the process of social identification determines how this reflects upon the *self*. This can either imply that the *self* is identified with that group and presumably shares its characteristic features, or lead to the conclusion that the *self* is distinct from that group and its features. Importantly, social identification not only refers to the *cognitive* awareness that one can be included in a particular group, but also incorporates the *emotional significance* of that group membership for the *self* (Tajfel, 1974, 1978a,

1978b). To the extent that people care about the groups they belong to (i.e., ingroups), they will be motivated to emphasize the distinct identity of those groups, and to uphold, protect, or enhance the value afforded to those groups and their members. On occasion (under conditions specified by the theory – see below), this may occur at the expense of other groups and their members (i.e., outgroups; Tajfel, 1978c).

Identity Management Strategies

SIT conceives of processes of social categorization, social comparison, and social identification as ways in which people actively define social reality and their own position relative to others in that reality (Tajfel, 1975, 1978b). The theory explicitly addresses the dynamic nature of social situations (Tajfel, 1974). Because the self is implicated in the group, people are motivated to emphasize and secure the ways in which their group is positively distinct from other groups (Tajfel, 1978c). Those who belong to what are generally seen as privileged groups (e.g., doctors, lawyers) should be motivated to enhance and retain their positive social identity. Clearly, though, many groups in society are devalued (e.g., the unemployed, migrants), and so the question of how members of these groups set about defining themselves positively becomes theoretically very important.

In this regard, a core feature of SIT is that it specifies different strategies that members of low-status social groups can adopt in order to address their situation and try to improve the value of their social identity. This in turn has profound implications for the ways in which members of high-status groups tend to protect and secure the current standing of their group (Tajfel, 1978c; Turner and Brown, 1978).

Individual mobility is an individual-level strategy whereby people may seek to escape, avoid, or deny belonging to a devalued group, and seek instead to be included in (or attempt to “pass” as a member of) a group of higher social standing (Tajfel, 1975). For instance, second-generation migrants may seek to escape their plight by pursuing an education or career that allows them to be seen as a member of a high-status professional group (e.g., as a lawyer) rather than as a member of an ethnic group that has low status in society. Individual mobility thus involves emphasizing how the *individual* self is different from other group members. But even if this helps improve

the status of specific group members, and furnishes them with a more positive social identity, individual mobility does not benefit (or even address) the standing of their ingroup as a whole.

[p. 383 ↓]

Social creativity refers to a process whereby group members seek to redefine the intergroup comparison by representing the ingroup in terms of positive rather than negative characteristics. This can be achieved in at least three ways: first, by focusing on other dimensions of intergroup comparison (e.g., comparing groups in terms of friendliness instead of material wealth); second, by including other groups in the comparison (e.g., as when migrants compare their economic success to those in their country of origin, rather than to others in the host society); and third, by changing the meaning of low-status group membership (e.g., as in the assertion that “black is beautiful”). Again, while this type of strategy is likely to help people cope with their devalued position in society, and may thereby benefit psychological wellbeing, it does not actually address or change the status quo or improve the ingroup's objective outcomes.

Social competition refers to a strategy whereby group members engage in forms of conflict designed to *change* the status quo (in ways that individual and social creativity do not). For instance, workers may seek to improve their work conditions or standard of living through union action, homosexuals may gain marriage or adoption rights due to political pressure, or women may improve their career prospects by pushing for the introduction of equal rights legislation. Social change can be contrasted with individual mobility in the sense that it explicitly addresses the situation of the group as a whole, where individual mobility seeks only to improve the social standing of particular individuals. Social change is also different from social creativity as it focuses on achieving changes to objective or material outcomes, whereas social creativity focuses primarily on a cognitive reinterpretation of the status quo. Importantly too, social competition involves concerted collective action oriented towards the achievement of change. Here groups compete with each other for superiority on a shared value dimension that reflects directly upon their mutual social standing.

Socio-Structural Characteristics

The final issue that is addressed in SIT concerns the conditions under which people are predicted to pursue these different strategies for social identity maintenance or improvement. Here the theory proposes that the way in which people respond to their group's circumstances depends on perceived characteristics of the prevailing social structure. Obviously, laws and cultural traditions or objective (im)possibilities may pose constraints on which forms of social identity improvement can be realistically achieved. However, the socio-structural characteristics to which SIT refers are explicitly defined as *subjective belief structures* regarding the opportunities ("cognitive alternatives" to the status quo) and valid motives for individual and group status improvement (Tajfel, 1975).

Permeability of group boundaries relates to the subjective belief that it is possible for individuals to act as independent agents within a given social system. Importantly, the main concern here is not whether it is possible to shed central or defining group characteristics, such as one's gender or ethnic origins. In cases as these, a full change of group membership is clearly not feasible. What matters in this context, is whether people feel that *by virtue of these defining group characteristics*, their access to other groups (and the material and psychological outcomes associated with them) is restricted, or whether they believe they can achieve a position in society that reflects their individual merit, regardless of their group membership. If they perceive group boundaries to be permeable they will be more inclined to pursue individual mobility as an attractive and viable strategy. On the other hand, if boundaries are perceived to be impermeable, then individuals are likely to feel more bound to their group. In this case, attempts at status improvement will tend to be pursued at the group level.

Stability of group status refers to the notion that some differences between groups are seen as fluid and as subject to change, [p. 384 ↓] while other differences tend to be regarded as more enduring and stable over time. To the extent that groups differ in concrete properties or abilities that are needed to achieve certain outcomes (e.g., physical endurance of male versus female athletes) differences between them are typically seen as inherent and unlikely to change. In many situations, however, people feel that differential group outcomes reflect historic developments (e.g., unequal access

to educational opportunities) or are the result of chance occurrences (e.g., differential numerical representation) rather than some essential or inherent difference in group value or deservingness. To the extent that status differences are thought to be stable, individuals with a devalued social identity are less likely to pursue strategies of social change and instead will be inclined to pursue a strategy individual mobility. However, if this is not possible due to impermeable group boundaries, then they should prefer to pursue social creativity strategies.

Legitimacy of current status relations refers to moral convictions that determine the *motivation* to change, where permeability and stability indicate perceived *opportunity* for change. Legitimacy can refer to a number of different aspects of a given social situation. In the first instance, the basis for including individuals in groups can be seen to be illegitimate in being based on incorrect assumptions about group-defining characteristics. This might arise, for instance, if a woman's gender rather than her professional qualifications were used to infer and ascribe her professional identity (e.g., so that she is treated as a secretary rather than as a manager). Status relations between groups can be also seen as illegitimate if important status-defining features (e.g., academic ability, professional competence) are selectively ascribed to some groups (e.g., men) rather than others under conditions where there is no objective indication that this is valid. Finally, the ascription of higher value to certain group characteristics can be seen as illegitimate. This might occur, for example, if the task-oriented leadership behaviors typically displayed by men are valued more highly than the socio-emotional leadership shown by women. Each of these forms of illegitimacy may motivate people to seek ways to rectify the current state of affairs. Importantly, this is not only the case for those who personally suffer from unjust treatment. Group members who benefit from unearned advantage may also be motivated to correct past injustice and support or sponsor the cause of those suffering from an unfair social system. This is the case, for instance, when senior white men in a company champion measures intended to improve career opportunities for young women or for members of minority ethnic groups.

Core Predictions

Building on the above principles, a number of core predictions were subsequently systematized within Tajfel and Turner's (1979) definitive statement of SIT (see also Tajfel and Turner, 1986). These predictions have received empirical support across a range of intergroup contexts and applied settings in studies using a number of different research methodologies (for a recent overview, see Haslam et al., 2010a).

- 1. To the extent that individuals internalize a group membership as a meaningful aspect of their self-concept, they will strive to make favorable comparisons between this group and relevant outgroups, in order to achieve or maintain a positive social identity.
- 2. As a result, social categorization can be sufficient to engender intergroup discrimination and intergroup conflict (i.e., in the absence of a conflict of interest over the division of resources or material outcomes, for example, as a result of historical antagonism).
- 3. The search for positive social identity may take different forms (*individual mobility, social creativity, social competition*), depending on consensual definitions of social reality that pertain to socially shared justifications (*legitimacy of group and individual outcomes*) and perceived cognitive alternatives to current status relations (*permeability of group boundaries and stability of status relations*).

[p. 385 ↓]

Theoretical Developments

After developing different elements of his thinking in the early 1970s (Tajfel, 1974, 1975), Tajfel contributed three chapters to the volume he edited in 1978 to more systematically describe the societal context and psychological mechanisms that inspired SIT (Tajfel, 1978a, 1978b, 1978c). These chapters can be seen to represent the intellectual origins of SIT. The further specification of its basic principles and mechanisms and how these impact on social behavior is laid down in the 1979

chapter (Tajfel and Turner, 1979). As suggested above, because this integrates and systematizes the theory's core ideas, this is typically cited as the definitive source of this theoretical perspective (at the time of writing this chapter has been cited more than 3,000 times). Another comprehensive overview of the central ideas underlying SIT and their implications for intergroup relations was provided by Henri Tajfel, in his 1982 *Annual Review of Psychology* chapter.

Over the years, however, several researchers have reported finding SIT ambiguous or unclear (e.g., Elsbach and Kramer, 1996; Jost and Elsbach, 2001; see also Brown, 2000). This in turn inspired others to clarify theoretical statements in an attempt to redress any misunderstandings, and to clarify exactly which predictions can (and cannot) be derived from the theory (e.g., Ellemers et al., 2003; Haslam and Ellemers, 2005; McGarty, 2001; Turner, 1985, 1999; Turner and Reynolds, 2001). As well as this, a number of researchers and different research groups have been involved in the development and extension of SIT in the process of testing and refining its core ideas.

Indeed, as a result of the large volume of research that it has inspired, various accounts of the theory can be found in the literature. These often emphasize specific aspects of the theory (e.g., the importance of positive distinctiveness) or focus on aspects of the theory relating to a specific concern (e.g., the determinants of collective action). Because of their variety, these developments can easily confuse those interested in learning about SIT, who – seeking an overview that addresses these different issues and connects the large body of recent empirical research to a single underlying framework – find it difficult to identify a core resource. While we would not claim the current contribution to constitute a comprehensive or definitive overview of developments relating to SIT, in the following we aim to summarize the main issues and concerns that have emerged over the years, to help explain how these both derive from and inform SIT.

Testing the Core Predictions

Early work in the 1970s was mainly concerned with development of the theory. Different researchers tested the validity of the mere categorization effects associated with the minimal group paradigm and the explanation provided to account for these effects.

Some of this work addressed questions that were raised concerning the methodology and measures that had been used in the original minimal group studies. This resulted in a series of studies that examined what happens when people are categorized according to different criteria (e.g., shared preferences, similar abilities, random assignment, Billig and Tajfel, 1973; Tajfel and Billig, 1974). It also led a body of research that explored different ways of assessing ingroup favoritism and which looked at the effects of using different outcome allocation methodologies (Bornstein et al., 1983; Turner, 1983).

An important conclusion from these efforts (see Bourhis et al., 1997; Diehl, 1990; Turner, 1999) was that the effects obtained in the minimal group studies could be reliably reproduced when methods specifically excluded the possibility that they arose from (a) material gains and instrumental benefit (as suggested by instrumental or economic approaches to intergroup relations), (b) a conflictual history (as suggested by sociological approaches to intergroup conflict), or (c) personality or *a priori* individual differences [p. 386 ↓] (as suggested by psychodynamic and other personological approaches).

However, one unintended consequence of these initial concerns and the discussions they generated was to evoke an impression that the minimal group studies and the effects they demonstrate constitute the essence of SIT. It is quite common, then, for secondary sources to suggest that the theory revolves around the idea that identification with a social group leads to ingroup bias, so that the ingroup is inevitably favored over any outgroup. Indeed, those who focus on the minimal group studies are often tempted to think that mere categorization is the main process, and that ingroup favoritism is the main outcome with which the theory is concerned. This is incorrect. Instead, the minimal group studies should be seen as an empirical demonstration of the importance of social identities for behavior that served as a catalyst for subsequent theorizing about nature and consequences of those identities. Thus, although the minimal group studies played a critical role in the conception and development of SIT, their significance lies in their historical status as a stimulus for the new way of thinking about the behavior of people in groups that SIT subsequently articulated, not in the fact that they capture the theory's core ideas.

The key problem with thinking that SIT is “all about” mere categorization and ingroup favoritism, is that this characterization (a) fails to recognize that the theory distinguishes

between a number of different identity-enhancement strategies (i.e., not only social competition but also individual mobility and social creativity), (b) neglects the fact that the core predictions of SIT refer to specific boundary conditions, and (c) overlooks the moderator variables that are predicted to impact on people's use of particular identity-enhancement strategies. Indeed, other studies that addressed the more complex and dynamic nature of intergroup relations in more messy or real-life situations, were in many ways far better suited to illustrate and test the core predictions made by SIT. These studies examined evidence for the pursuit of different identity management strategies under specific conditions (e.g., Ellemers, 1993; Lalonde and Silverman, 1994; Reicher and Haslam, 2006). And, in line with SIT's predictions, they found that depending on relevant socio-structural conditions people pursue strategies of either individual mobility, social creativity, or social competition (for overviews see Bettencourt et al., 2001; Brown, 2000).

Interdependence and Bias

Following on from early discussion of the minimal group findings, some of the work in the 1980s still sought to compare the effects of realistic conflict and outcome interdependence with those that arose from mere categorization (e.g., Rabbie et al., 1989). In a way, this is an issue that can never be resolved, as this work has convincingly shown that mere categorization effects as well as interdependence and conflict can each in and of themselves promote displays of ingroup favoritism. In another way, though, this demonstration also indicates that this debate *has* been resolved. For the fact that outcome interdependence can promote ingroup favoritism and intensify intergroup conflict does not make it less interesting that (as SIT suggests) mere categorization has similar effects, even in the *absence* of such alternative or additional concerns. Indeed, while the minimal group studies demonstrated that social categorization can be *sufficient* to raise ingroup favoritism, it was never argued that categorization is *necessary* for such effects to emerge. Thus, over the years, it has come to be generally understood that SIT complements realistic conflict theory and that social identity concerns can interact with instrumental concerns. At the same time it is acknowledged that SIT was not intended, and should not be seen, to displace an interest in realistic causes of intergroup conflict (Brown, 2000).

[p. 387 ↓]

Self-Esteem

During the 1990s the emphasis in social identity research shifted from demonstrating the basic phenomenon of intergroup differentiation, to examining *why* people might be motivated to act in ways that reflected a group-level definition of self rather than an individual-level definition. These efforts related to the so-called “self-esteem hypothesis,” advanced by Hogg and Abrams. This proposed two corollaries of SIT’s core predictions: (a) successful intergroup discrimination should elevate self-esteem, and (b) depressed or threatened self-esteem should promote intergroup discrimination (Abrams and Hogg, 1988; Hogg and Abrams, 1990). Research then examined the role of self-esteem both as a cause and as an effect of positive intergroup differentiation.

Unfortunately, much of this work was plagued by conceptual and methodological imprecisions, leading to empirical studies that generated a range of seemingly inconsistent results (for overviews, see Brown, 2000; Long and Spears, 1997; Turner, 1999). For instance, personal-level or enduring self-esteem measures were used to assess shifts as a result of intergroup differentiation. Arguably, however, the predicted effects should emerge on more situationally defined measures that tap into group-based and collective-level aspects of self-esteem (Crocker and Luhtanen, 1990; Rubin and Hewstone, 1998). In fact, it was noted that individual- and group-level processes may even interact in that those who hold high levels of personal self-esteem should be most inclined to defend and uphold the status of their group when it is devalued by others (Long and Spears, 1997). Furthermore, it was pointed out that the self-esteem hypothesis focuses on intergroup differentiation as the primary response to social identity threat, while failing to acknowledge that alternative (individual-level as well as group-level) strategies may be used to (re)establish a positive social identity (Turner, 1999). Thus, although research efforts only found mixed support for the self-esteem hypothesis, over the years different scholars noted that both the formulation of these corollaries and the way in which they had been tested were flawed (Long and Spears, 1997; Rubin and Hewstone, 1998; see also Turner, 1999).

The self-esteem hypothesis was also criticized for more metatheoretical reasons. In particular, Turner (1999) voiced a concern that to focus on self-esteem as the critical factor in intergroup differentiation could lead to a misrepresentation of the theory's core ideas. For if positive ingroup differentiation is seen to be driven by self-esteem needs, then this may be taken as implying that, in situations where the individual self is connected to others in a group, intergroup behaviors primarily serve individual-level motives (e.g., associated with outcome interdependence). This emphasis on individual-level instrumental needs and concerns is at odds with the group-level approach that is central to the SIT (and metatheory). Specifically, it goes against the theory's concern to provide a cognitive and psychological account of social perception and behavior that goes beyond individual-level rational choices or cost–benefit analysis.

The lack of evidence for a simple relation between self-esteem and intergroup differentiation also led to the proposal of alternative motives and additional factors that might determine when, how, and why intergroup differentiation is expressed. Reduction of uncertainty about the position of the self in relation to others was proposed as a broader motive that may induce identity enhancement as well as other responses (Hogg and Abrams, 1993; Hogg and Mullin, 1999). Furthermore, researchers demonstrated the force of social reality constraints in determining how and when people express their internal convictions about the worth of their group (see Ellemers et al., 1999a, for an overview).

Importantly, while these developments can be seen as further specifications of SIT and its predictions, they do not really propose novel or contradictory ideas (instead, they essentially flesh out one aspect of the theory's original formulation). That is, the motive of [p. 388 ↓] uncertainty reduction is fully consistent with the original idea that people use social categorizations as a basis for imbuing novel situations with meaning. Likewise, the notion that members of consensually devalued groups may refrain from making public claims of ingroup superiority is in line with the original idea that existing power or status differences between groups determine which identity management strategies are likely to prove feasible.

Self-Categorization

During the 1980s and early 1990s, Turner and his colleagues (Turner, 1982, 1985; Turner et al., 1987, 1994) set out to elaborate on the cognitive processes that underpin group-rather than individual-level conceptions of self and others. These developments led to the formulation of *self-categorization theory* (SCT). This theory further specifies and extends Tajfel's original proposition that social categorization serves as a basis for understanding and responding meaningfully to complex social situations (Tajfel, 1969). However, SCT focuses more explicitly on the fact that social categorizations can be made at different (nested) levels of inclusiveness or abstraction (e.g., Londoner, UK citizen, European) and that the same individual can be included in multiple categories on the basis of different (cross-cutting) criteria (e.g., as a woman, as a German, as a psychologist). Within SCT, these ideas were formalized in terms of a number of core assumptions and related hypotheses (Turner, 1985). In particular, these asserted that:

- 1. The self is represented cognitively in terms of self-categories that can be defined at different levels of abstraction. These range from exclusive self-categorization in terms of personal identity (e.g., "I, Christine") to inclusive self-categorization in terms of broad social identities (e.g., "us Dutch").
- 2. The formation of self-categories is partly a function of the metacontrast between interclass and intraclass differences. This means that people will tend to define themselves in terms of a particular self-category (e.g., as Dutch) to the extent that the differences between members of that category on a given dimension of judgment are perceived to be smaller than the differences between members of that category and others that are salient in a particular context (e.g., Belgians, Germans).
- 3. Metacontrast also partly determines the internal structure of self-categories and the prototypicality of particular category exemplars. This means that a person's capacity to represent and embody a given social category increases to the extent that the differences between them and other members of that category are smaller than the differences between them and members of other categories that are salient in a particular context.
- 4. The salience of a particular self-category leads to the accentuation of perceived intraclass similarities and interclass differences. In this way,

patterns of assimilation and contrast reflect the relative inter-changeability of category exemplars in relation to a currently salient self-categorization.

Additional mechanisms of category accessibility and normative fit were also specified in order to explain and predict which self-categorization is most likely to be used in a given situation (Oakes, 1987; Turner, 1985). In explaining how individuals come to define themselves in terms of one social identity rather than another, SCT thus emphasizes the importance of a range of contextual contexts that elements contexts that serve to make one particular social self-categorization more meaningful than others. The principles it articulates also predict that the same objective group membership will be experienced differently, depending on the groups with which an ingroup is compared and the situation in which these comparisons are made (Haslam and Turner, 1992).

Importantly too, SCT introduced the concept of *depersonalization*, to describe the psychological process through which people come to perceive the self as an interchangeable exemplar of a social category, rather than as a separate individual with unique traits (i.e., so that the self is defined in terms of social identity rather than personal [p. 389 ↓] identity; Turner, 1982). Moreover, the theory also specifies the behavioral consequences of depersonalization, arguing that it is this process that makes group behavior possible. In particular, Turner (1982) hypothesized, and early studies confirmed, that depersonalization is a basis for group cohesion, interpersonal attraction, and social cooperation (e.g., Hogg, 1992; Hogg and Turner, 1985). Later work also elaborated upon the implications of self-categorization and depersonalization for processes of stereotyping (Oakes and Turner, 1990; Oakes et al., 1994), social influence (Turner, 1991), and leadership (Turner and Haslam, 2001). More recently still, research has shown that self-categorization processes play a key role in the expression of personal identity (Turner and Onorato, 1999; Turner et al., 2006).

In sum, building upon SIT's original concern with processes of social categorization and its clarification of the importance of social identity for intergroup behavior, SCT provides a more detailed and more general account of the psychological mechanisms that lead individuals to define themselves in terms of particular group memberships and to act in terms of those group memberships. Although they have somewhat different foci and contain a number of quite different hypotheses, SIT and SCT thus share a range of key assumptions, particularly in metatheoretical terms. Accordingly, in recent publications,

the two theories are often presented as complementary theoretical frameworks, and treated as component parts of an integrated *social identity approach* or perspective (e.g., Reicher et al., 2010; Turner and Reynolds, 2001).

Conceptualizations and Measures

Around the turn of the century it became clear that, as it had evolved, empirical research had incorporated a number of different conceptualizations and measures of social identity as a construct. Accordingly, a number of researchers set out to compare these different conceptualizations and to examine their theoretical and empirical implications. A first source of confusion in this respect was that some researchers used the term “social identity” to refer to the *content* of characteristics typically associated with a particular social group (i.e., social identity *value*), while others used the same term to indicate the extent to which an individual subjectively perceived the self to be *included* in the group (i.e., social identity *strength*). To some extent increased awareness of this potential confusion served to resolve it, as more careful and consistent use of terminology (e.g., positive/negative social identity versus level of social identification) certainly helps to clarify what is meant by any given reference.

Nevertheless, this more specific use of terminology also raised further conceptual questions about the “essence” or “true” definition of social identity (e.g., Kreiner and Ashforth, 2004; Mael and Tetrick, 1992; Van Dick, 2004). Do people’s social identities have such profound implications because of the cognitive inclusion of the self in the group, or because of their experience of affective commitment to the group? It is probably no surprise that this question became important, given the many different ways in which researchers had been drawing on social identity and SCTs. This contributed to the view that the (cognitive) process of self-categorization (the focus of SCT) should be seen as distinct from (subjective feelings of) social identification (addressed in SIT).

To some extent this difference in emphasis simply reflects the difference in the two theories’ explanatory aims and foci, as indicated above. Importantly, though, a careful reading of the original theoretical formulations reveals that *both* SCT and SIT address the way in which cognitive and affective components of social identity are *connected*, as the cognitive awareness of a certain category membership is a

necessary precondition that has to be met before one's group membership can acquire any emotional significance. In line with this point, the original definition quoted above [p. 390 ↓] (Tajfel, 1974) makes it clear that the knowledge and emotional significance of group membership *together* comprise social identity. Nevertheless, because this conceptualization incorporates these different aspects, they can be considered either as separate components or as comprising a single overarching construct. Accordingly, it has been convincingly argued, and shown, that – depending on the issue under concern – it can be useful to focus either on specific components of social identity or to address social identity as a broader multidimensional construct (Ellemers et al., 1999b; Hogg, 1992; Leach et al., 2008; Ouwerkerk et al., 1999).

Treating identification as a multidimensional construct also made it possible to further specify different forms of identity threat that ensue when different identity components are not aligned (Branscombe and Doosje, 2004; Branscombe et al., 1999a; Doosje et al., 1989). For instance, people may experience social identity threat when they feel emotionally involved with a group in which they are not cognitively included, or vice versa. This is likely to happen, for instance, in the process of transitioning to another group during individual mobility, or when internal definitions of self do not correspond to the way in which one is treated by others. Additionally, this work has made it clear that both cognitive and emotional components of identification have to exceed a minimum threshold level before individuals can be expected to respond in terms of their group-based identity. That is, while those who are cognitively included but not emotionally involved in the group may act as group members when individual-level concerns make it attractive for them to do so (e.g., for self-presentational reasons, or for fear of social sanctions), people who are both cognitively included and emotionally involved in the group tend to define the self as group members more consistently across different multiple situations and contexts (Barreto and Ellemers, 2000; Ellemers et al., 2002).

Treating social identity as an important source of behavioral motivation also raised the issue of how it should be conceptualized in relation to other theoretical models of motivation. Over the years, this meant that a number of different questions were raised by researchers from inside and outside the social identity tradition. Should social identity be seen as an individual difference variable indicating chronic levels of altruism or empathy? Does it indicate different psychological needs stemming from early attachment experiences or cultural variations in how we think of individuals and groups

(reflecting levels of individualism versus collectivism)? Does it reflect interpersonal or situational variations in a more generic need to belong? Should it primarily be seen as a (stable) cause of intergroup behavior across different contexts, or as a (situational) effect of specific intergroup comparisons?

Empirical studies that addressed these issues generated an abundance of evidence that the effects of social identification are tied to *specific* groups. Hence, group identification does not simply indicate “need to belong,” as people prefer to be included and valued by specific groups rather than by groups in general (e.g., Ellemers et al., 2004b). Likewise, the level of identification with the same group tends to change over time and across social contexts (e.g., Doosje et al., 2002). This speaks against a more stable individual difference or cultural difference approach, and is consistent with SIT’s core notion that group identification indicates a person’s situational inclination to think and act in terms of a group-level self (i.e., as “us” rather than “I”). Indeed, levels of group identification have been shown to develop and change as part of a recursive process (Branscombe et al., 1999b; Schmitt and Branscombe, 2002) in which initial levels of identification determine how people respond to intergroup situations, and this response in turn intensifies or diminishes feelings of identification. Thus, even if researchers have sought to understand this complex process by administering measures [p. 391 ↓] of social identification and addressing simple causal relations at specific points of the recursive cycle (i.e., treating identification either as a cause or as an effect of other responses), it is important to bear in mind that identification was originally defined as a dynamic construct and needs always to be understood in these terms.

Intellectual Impact of Ideas

SIT can be characterized as a “grand theory.” It addresses intrapersonal cognitive mechanisms, interpersonal and intergroup behaviors, and social relations, and connects processes occurring at these different levels of analysis to provide a broad theoretical framework that can help understand a range of phenomena. In our review of the theory, we have tried to convey this complex nature of the theory, and have shown how, after its initial formulation, different aspects were further developed and specified. In this sense, then, the theory is not a fixed number of statements, but represents more of

a “living” creature to which different theorists and researchers have each added their embellishments, additions, and refinements over the years.

The broad nature of SIT at the same time can be seen both as a strength and as a weakness. On the one hand, the number of references to SIT in empirical studies clearly attests to its explanatory power, and shows that it can be applied to a range of issues. On the other hand, the fact that the theory cannot be distilled into a simple mantra or summarized in a limited number of simple hypotheses that “always work” (without consideration being given to localized context), implies that the broader constellation of ideas is not easy to test or refute. Indeed, the emphasis on dynamic changes and contextual differences, and the realization that different response patterns may emerge depending on specific circumstances, easily evokes the impression that SIT can be used to explain anything and everything – but only after the fact. At the same time, the pressure to reduce to the theory to a simple message, (e.g., of the form “social identification leads to intergroup discrimination”) has contributed to vulgarized textbook summaries that do violence to the theory’s core character.

What then is the added value of SIT? We think the answer to this question lies at least in part in this realization that it is different from many other theories in social psychology. Rather than consisting of a limited set of specific predictions, it represents a particular metatheoretical approach, that provides a unique *perspective* on social cognition and social behavior. That is, the general notion that in addition to the personal level of self-definition people may also self-define at the group level (and can switch between these different levels), helps make sense of a range of phenomena that prove hard to explain in terms only of individual-level psychological mechanisms. Accordingly, SIT opens up the possibility of considering whether a group-level approach can help understand a particular phenomenon, and, if this is the case, provides conceptual tools that can usefully inform and structure this type of analysis.

Importantly, this is not to say that each and every issue in social psychology should be addressed as a group-level problem. If only for reasons of theoretical parsimony, more basic or individual-level explanations should be favored when appropriate. Nevertheless, an important driver for the development and application of SIT was an awareness that researchers are generally far more likely to fall into the opposite trap and overlook the appropriateness of group-level analysis. As Turner and Oakes (1997)

observe, social psychology (and psychology more generally) has not suffered from an underemphasis on the psychology of the individual, or a shortage of individual-level theories. Thus, if anything, there seems to be a general tendency in social psychology to explain anything and everything from individual-level mechanisms, thereby neglecting the power of the group to inform and motivate social behavior, and excluding the possibility of the group-level self. A key contribution of SIT [p. 392 ↓] has been to redress this balance, and provide researchers with analytical options that have otherwise been lacking.

Social Issues and Applications

Even though the original social categorization studies were very detached and artificial, the theory these gave rise to was focused explicitly on the task of analyzing and explaining social relations in the world at large. When it was developed, SIT brought together insights from a number of intellectual and research traditions, but nevertheless provided an important new perspective and focus on intergroup relations that was genuinely revolutionary. By examining how specific characteristics of social contexts interact with individual cognitive processes, and explaining the origins and consequences of a conceptualization of self at the group level, the theory spells out a number of important ideas about social psychological functioning that have relevance to a host of situations.

In view of Tajfel's original goal of seeking to understand the emergence of conflict in intergroup relations, SIT's original formulations focused closely on social behavior in situations defined by historical between-group differences in power and status. Nevertheless, the prominence of the minimal group studies and SCT's examination of basic cognitive processes makes it easy to overlook the fact that SIT has often been used to examine interactions occurring between members of real social groups. Over the years, then, researchers have used the theory to help them understand tensions between ethnic, religious, or linguistic groups, and to examine and predict responses to migration, changing labor relations, and the development of group motivation. This work both informed and helped develop SIT as researchers examining various group types under different conditions became aware of specific complexities, moderating variables, and boundary conditions that were relevant to the theory's core predictions.

As such work accumulated, it served to validate SIT's core predictions regarding the conditions under which people would pursue particular self-enhancement strategies and experience different forms of identity threat. This support was found across different types of intergroup comparisons and for various sources of group value (Mullen et al., 1992), such as power, status (Sachdev and Bourhis, 1991), or group size (Simon and Brown, 1987). However, the existence of these real and objective differences between groups in social contexts also made it clear that it is not always realistic (or desirable) for group members to strive to make positive intergroup comparisons. When the achievement of positive intergroup comparisons is not feasible or would unduly antagonize an outgroup, group members may seek distinctiveness from other groups (Mummendey and Schreiber, 1983, 1984), especially when the differences between them are ambiguous or ill-defined (Jetten and Spears, 2004; Jetten et al., 2004). For similar reasons, maintenance of current intergroup distinctions may be preferred over attempts to enhance or improve one's social identity (Ellemers et al., 1992; Scheepers and Ellemers, 2005).

While the theory addresses each level of analysis in turn to analyze and understand the psychological mechanisms relevant to individual- and group-level behavior, research has simultaneously considered the effects of intra- and inter-group comparisons. This work has indicated that an awareness of intragroup heterogeneity and individuality does not necessarily exclude the formation of a common group identity (Doosje et al., 1999; Hornsey and Jetten, 2004; Postmes and Jetten, 2006; Rink and Ellemers, 2007; Simon, 1992), and that a positive social identity depends as much on evaluations of the self by others in the group as on evaluations of the group by other groups (Branscombe et al., 2002; Smith et al., 2003; Tyler and Blader, 2000). Finally, the examination of more complex and rich intergroup situations has made it clear that people do not always engage with the intergroup comparisons that [p. 393 ↓] others invite them to make, but actively define and carve out their social identity from multiple dimensions (Derkx et al., 2007), sources of group value (Leach et al., 2007), and group identities (Spears and Manstead, 1989) available to them in real life.

While all of these insights represent important extensions of social identity theorizing, all remain consistent with the theory's core premises. Moreover, the "grand" nature of SIT and its explicit consideration of social contextual variables in addition to individual-level cognitive processes, makes obvious its relevance to a broad variety of social problems

and issues in organizational behavior (see also Ashforth and Mael, 1989; Haslam, 2004; Haslam and Ellemers, 2005; Haslam et al., 2003; Hogg and Terry, 2000). In this vein, most recently, insights from SIT have been used to examine *individual* wellbeing and performance, documenting the implications of people's social identities for the experience of stress, work outcomes, and physical and mental health (e.g., Haslam and Reicher, 2006; Haslam et al., 2009; Scheepers and Ellemers, 2005).

SIT has also proved useful when it comes to understanding *interpersonal* behavioral alignment, for instance when participating in political activities or social protest (Reicher, 1987; Simon and Klandermans, 2001; Wright, 2000). Along related lines, the theory has also been used to analyze and improve *intergroup relations*, for instance by considering interethnic conflict or gender discrimination in terms of social identity concerns (Ellemers et al., 2004c; Ryan and Haslam, 2007).

Intertwined with much of this work, SCT has also been used to further our understanding of important *group dynamics*, particularly those relating to social influence and group polarization (e.g., Levine et al., 2000, 2005; Postmes et al., 2005; Smith et al., 2003; Turner, 1991; Wetherell, 1987). In organizational contexts, this has also led to important insights into processes of leadership (Haslam et al., 2010b; Hogg and Van Knippenberg, 2004; Reicher et al., 2005; Turner and Haslam, 2001), communication (Postmes, 2003), and work motivation and group performance (Ellemers et al., 2004a).

Conclusion

In this chapter we have outlined the core ideas and basic premises of SIT. Supported by a large body of empirical evidence, this theory argues for the importance of distinguishing between social psychological processes at individual, interpersonal, group, and intergroup levels. We have noted that over the years, different accounts or selected ideas have been seen as representing the essence of the theory, resulting in a number of controversies, updates, refinements, and expansions. For some, this may create an impression that theory is overly complex and controversial, and hence not particularly useful as an analytical framework. Against this conclusion, however, it is apparent that SIT has inspired, and been supported by, a large body of important

empirical studies and has informed a range of important theoretical developments in social psychology and cognate disciplines. This work makes it clear that the theory provides an analysis of complex social phenomena that can help researchers understand and address a number of important social issues and problems.

Importantly too, because SIT is addressed to the process of *social change*, it also points to the fact that social psychological processes do not simply contribute to the reproduction of the status quo, but also help to bring about change in the world. In this sense too, the theory is progressive and optimistic, rather than conservative and pessimistic. Instead of being reductionist and deterministic, it offers scope for interventions that can help improve individual wellbeing, group interactions, and social relations.

Thus, SIT is more than a metaphor: it provides a different way of thinking about individuals and groups, with an explicit emphasis on the impact of social contextual factors. Even though this makes the theory “grand” [p. 394 ↓] and complex, a careful consideration of its core ideas makes it clear that a relatively limited set of ideas and variables can help understand a range of phenomena across different situations and settings. In order to appreciate this point, there is much to be gained by going back to Tajfel and Turner’s original writings (recently reprinted in Postmes and Branscombe, 2010) and enjoying their ideas and lucid explanations firsthand.

NaomiEllemers S. AlexanderHaslam

References

Abrams, D. and Hogg, M.A. Comments on the motivational status of self-esteem in social identity and intergroup discrimination . European Journal of Social Psychology , 18, 317–334. (1988) <http://dx.doi.org/10.1002/ejsp.2420180403>

Ashforth, B. and Mael, F. Social identity theory and the organization . Academy of Management Review , 14, 20–39. (1989)

Barreto, M. and Ellemers, N. You can't always do what you want: Social identity and self-presentational determinants of the choice to work for a low status group . *Personality and Social Psychology Bulletin* , 26, 891–906. (2000) <http://dx.doi.org/10.1177/01461672002610001>

Bettencourt, B.A., Charlton, K., Dorr, N. and Hume, D.L. Status differences and in-group bias: A meta-analytic examination of the effects of status stability, status legitimacy, and group permeability . *Psychological Bulletin* , 127, 520–542. (2001) <http://dx.doi.org/10.1037/0033-2909.127.4.520>

Billig, M. and Tajfel, H. Social categorization and similarity in intergroup behavior . *European Journal of Social Psychology* , 3, 27–52. (1973) <http://dx.doi.org/10.1002/ejsp.2420030103>

Bornstein, G., Crum, L., Wittenbraker, J., Harring, K., Insko, C.A. and Thibaut, J. On the measurement of social orientations in the minimal group paradigm . *European Journal of Social Psychology* , 13, 321–350. (1983) <http://dx.doi.org/10.1002/ejsp.2420130402>

Bourhis, R.Y., Turner, J.C. and Gagnon, A. (1997) Interdependence, social identity and discrimination . In R. Spears, ed. , P.J. Oakes, ed. , N. Ellemers, ed. and S.A. Haslam (eds), *The Social Psychology of Stereotyping and Group Life* , pp. 273–295. Oxford: Blackwell.

Branscombe, N.R., ed. and Doosje, B. (eds) (2004) *Collective Guilt: International Perspectives* . Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139106931>

Branscombe, N., Ellemers, N., Spears, R. and Doosje, B. (1999a) The context and content of identity threat . In N. Ellemers, ed. , R. Spears, ed. , and B. Doosje (eds), *Social Identity: Context, Commitment, Content* , pp. 35–58. Oxford: Blackwell.

Branscombe, N.R., Schmitt, M.T. and Harvey, R.D. Perceiving pervasive discrimination among African-Americans: Implications for group identification and well-being . *Journal of Personality and Social Psychology* , 77, 135–149. (1999b) <http://dx.doi.org/10.1037/0022-3514.77.1.135>

Branscombe, N., Spears, R., Ellemers, N. and Doosje, B. Effects of intragroup and intergroup evaluations on group behavior . *Personality and Social Psychology Bulletin* , 28, 744–753. (2002) <http://dx.doi.org/10.1177/0146167202289004>

Brown, R. Social identity theory: past achievements, current problems and future challenges . *European Journal of Social Psychology* , 30, 745–778. (2000) <http://dx.doi.org/10.1002/1099-0992%28200011/12%2930:6%3C745::AID-EJSP24%3E3.0.CO;2-O>

Crocker, J. and Luhtanen, R. Collective self-esteem and ingroup bias . *Journal of Personality and Social Psychology* , 58, 60–67. (1990) <http://dx.doi.org/10.1037/0022-3514.58.1.60>

Derks, B., Van Laar, C. and Ellemers, N. Social creativity strikes back: Improving low status group members' motivation and performance by valuing ingroup dimensions . *European Journal of Social Psychology* , 37, 470–493. (2007) <http://dx.doi.org/10.1002/ejsp.375>

Diehl, M. The minimal group paradigm: Theoretical explanations and empirical findings . *European Review of Social Psychology* , 1, 263–292. (1990) <http://dx.doi.org/10.1080/14792779108401864>

Doosje, B., Branscombe, N.R., Spears, R. and Manstead, A.S.R. Guilty by association: When one's group has a negative history . *Journal of Personality and Social Psychology* , 75, 872–886. (1989) <http://dx.doi.org/10.1037/0022-3514.75.4.872>

Doosje, B., Spears, R. and Ellemers, N. Social identity as both cause and effect: The development of group identification in response to anticipated and actual changes in the intergroup status hierarchy . *British Journal of Social Psychology* , 41, 57–76. (2002) <http://dx.doi.org/10.1348/014466602165054>

Doosje, B., Spears, R., Ellemers, N. and Koomen, W. Group variability in intergroup relations: The distinctive role of social identity . *European Review of Social Psychology* , 10, 41–74. (1999) <http://dx.doi.org/10.1080/14792779943000017>

Ellemers, N. The influence of socio-structural variables on identity enhancement strategies . European Review of Social Psychology , 4, 27–57. (1993) <http://dx.doi.org/10.1080/14792779343000013>

Ellemers, N., Barreto, M. and Spears, R. (1999a) Commitment and strategic responses to social context . In N. Ellemers, ed. , R. Spears, ed. , and B. Doosje (eds), Social Identity: Context, Commitment, Content , pp. 127–146. Oxford: Blackwell.

Ellemers, N., De Gilder, D. and Haslam, S.A. Motivating individuals and groups at work: A social identity perspective on leadership and group performance . Academy of Management Review , 29, 459–478. (2004a)

Ellemers, N., Doosje, B. and Spears, R. Sources of respect: The effects of being liked by ingroups and outgroups . European Journal of Social Psychology , 34, 155–172. (2004b) <http://dx.doi.org/10.1002/ejsp.196>

Ellemers, N., Doosje, E.J., Van Knippenberg, A. and Wilke, H. Status protection in high status minorities . European Journal of Social Psychology , 22, 123–140. (1992) <http://dx.doi.org/10.1002/ejsp.2420220203>

Ellemers, N., Kortekaas, P. and Ouwerkerk, J. Self-categorization, commitment to the group and group self-esteem as related but distinct aspects of social identity . European Journal of Social Psychology , 29, 371–389. (1999b) <http://dx.doi.org/10.1002/%28SICI%291099-0992%28199903/05%2929:2/3%3C371::AID-EJSP932%3E3.0.CO;2-U>

Ellemers, N., Platow, M., Van Knippenberg, D. and Haslam, A. (2003) Social identity at work: Definitions, debates, and directions . In A. Haslam, ed. , D. Van Knippenberg, ed. , M. Platow, ed. and N. Ellemers (eds), Social Identity at Work: Developing Theory for Organizational Practice , pp. 3–28. London: Psychology Press.

Ellemers, N., Spears, R. and Doosje, B. Self and social identity . Annual Review of Psychology , 53, 161–186. (2002) <http://dx.doi.org/10.1146/annurev.psych.53.100901.135228>

Ellemers, N., Van den Heuvel, H., De Gilder, D., Maass, A. and Bonvini, A. The underrepresentation of women in science: Differential commitment or the Queen-

bee syndrome? *British Journal of Social Psychology* , 43, 315–338. (2004c) <http://dx.doi.org/10.1348/0144666042037999>

Elsbach, K.D. and Kramer, R.D. Members' responses to organizational identity threats: Encountering and countering the Business Week rankings . *Administrative Science Quarterly* , 41, 442–476. (1996) <http://dx.doi.org/10.2307/2393938>

Festinger, L. A theory of social comparison processes . *Human Relations* , 7, 117–140. (1954) <http://dx.doi.org/10.1177/001872675400700202>

Haslam, S.A. (2004) *Psychology in Organizations: The Social Identity Approach* , 2nd Edition. London: Sage.

Haslam, S.A. and Ellemers, N. (2005) Social identity in industrial and organizational psychology: Concepts, controversies and contributions . In G.P. Hodgkinson, ed. and J.K. Ford (eds), *International Review of Industrial and Organizational Psychology* , 20, 39–118. Chichester: Wiley.

Haslam, S.A., Ellemers, N., Reicher, S., Reynolds, K.J. and Schmitt, M.T. (2010a) The social identity perspective: An assessment of the impact and trajectory of its defining ideas . In T. Postmes, ed. and N. Branscombe (eds), *Rediscovering Social Identity: Core Sources* . New York: Psychology Press.

Haslam, S.A., Jetten, J., Postmes, T. and Haslam, C. Social identity, health and well-being: An emerging agenda for applied psychology . *Applied Psychology: An International Review* , 58, 1–23. (2009) <http://dx.doi.org/10.1111/j.1464-0597.2008.00379.x>

Haslam, S.A. and Reicher, S.D. Stressing the group: Social identity and the unfolding dynamics of stress . *Journal of Applied Psychology* , 91, 1037–1052. (2006) <http://dx.doi.org/10.1037/0021-9010.91.5.1037>

Haslam, S.A., Reicher, S.D. and Platow, M.J. (2010b) *The New Psychology of Leadership: Identity, Influence and Power* . London: Psychology Press.

Haslam, S.A. and Turner, J.C. Context-dependent variation in social stereotyping 2: The relationship between frame of reference, self-categorization and accentuation . European Journal of Social Psychology , 22, 251–278. (1992) <http://dx.doi.org/10.1002/ejsp.2420220305>

Haslam, S.A., ed. , Van Knippenberg, D., ed. , Platow, M., ed. and Ellemers, N. (eds) (2003) Social Identity at Work: Developing Theory for Organizational Practice . New York: Psychology Press.

Hogg, M.A. (1992) The Social Psychology of Group Cohesiveness: From Attraction to Social Identity . Hemel Hempstead: Harvester Wheatsheaf.

Hogg, M.A. and Abrams, D.A. (1990) Social motivation, self-esteem and social identity . In D. Abrams, ed. , and M.A. Hogg (eds), Social Identity Theory: Constructive and Critical Advances , pp. 28–74. London: Harvester Wheatsheaf.

Hogg, M.A. and Abrams, D.A. (1993) Towards a single-process uncertainty-reduction model of social motivation in groups . In M.A. Hogg, ed. and D. Abrams (eds), Group Motivation: Social Psychological Perspectives , pp. 173–190. New York: Harvester Wheatsheaf.

Hogg, M.A. and Mullin, B.A. (1999) Joining groups to reduce uncertainty: Subjective uncertainty reduction and group identification . In D. Abrams, ed. and M.A. Hogg (eds), Social Identity and Social Cognition , pp. 249–279. Oxford: Blackwell.

Hogg, M.A. and Terry, D.J. Social identity and self-categorization processes in intergroup contexts . Academy of Management Review , 25, 121–140. (2000)

Hogg, M.A. and Turner, J.C. Interpersonal attraction, social identification and psychological group formation . European Journal of Social Psychology , 15, 51–66. (1985) <http://dx.doi.org/10.1002/ejsp.2420150105>

Hogg, M.A. and Van Knippenberg, D. Social identity and leadership processes in groups . Advances in Experimental Social Psychology , 35, 1–52. (2004) <http://dx.doi.org/10.1016/S0065-2601%2803%2901001-3>

Hornsey, M.J. and Jetten, J. The individual within the group: Balancing the need to belong with the need to be different . *Personality and Social Psychology Review* , 8, 248–264. (2004) http://dx.doi.org/10.1207/s15327957pspr0803_2

Jetten, J. and Spears, R. The divisive potential of differences and similarities: The role of intergroup distinctiveness in intergroup differentiation . *European Review of Social Psychology* , 14, 203–241. (2004) <http://dx.doi.org/10.1080/10463280340000063>

Jetten, J., Spears, R. and Postmes, T. Intergroup distinctiveness and differentiation: A meta-analytical investigation . *Journal of Personality and Social Psychology* , 86, 862–879. (2004) <http://dx.doi.org/10.1037/0022-3514.86.6.862>

Jost, J.T. and Elsbach, K.D. (2001) How status and power differences erode personal and social identities at work: A system justification critique of organizational applications of social identity theory . In M.A. Hogg, ed. and D.J. Terry (eds), *Social Identity Processes in Organizational Contexts* , pp. 181–196. Philadelphia: Psychology Press.

Kreiner, G.E. and Ashforth, B.E. Evidence toward an expanded model of organizational identification . *Journal of Organizational Behavior* , 25, 1–27. (2004) <http://dx.doi.org/10.1002/job.234>

Lalonde, R.N. and Silverman, R.A. Behavioral preferences in response to social injustice: The effects of group permeability and social identity salience . *Journal of Personality and Social Psychology* , 66, 78–85. (1994) <http://dx.doi.org/10.1037/0022-3514.66.1.78>

Leach, C., Ellemers, N. and Barreto, M. Group virtue: The importance of morality vs. competence and sociability in the evaluation of in-groups . *Journal of Personality and Social Psychology* , 93, 234–249. (2007) <http://dx.doi.org/10.1037/0022-3514.93.2.234>

Leach, C.W., Van Zomeren, M., Zebel, S., Vliek, M.L.W., Pennekamp, S.F., Doosje, B., Ouwerkerk, J.W. and Spears, R. Group-level self-definition and self-investment: A hierarchical (multi-component) model of in-group identification . *Journal of Personality and Social Psychology* , 95, 144–165. (2008) <http://dx.doi.org/10.1037/0022-3514.95.1.144>

Levine, J.M., Higgins, T.E. and Choi, H.S. Development of strategic norms in groups . Organizational Behavior and Human Decision Processes , 82, 88–101. (2000) <http://dx.doi.org/10.1006/obhd.2000.2889>

Levine, J.M. and Moreland, R.L. (1987) Social comparison and outcome evaluation in group contexts . In J.C. Masters, ed. and W.P. Smith (eds), Social Comparison, Social Justice, and Relative Deprivation: Theoretical, Empirical and Policy Perspectives , pp. 105–127. Hillsdale, NJ: Lawrence Erlbaum.

Levine, R.M., Prosser, A., Evans, D. and Reicher, S.D. Identity and emergency intervention: How social group membership and inclusiveness of group boundaries shapes helping behavior . Personality and Social Psychology Bulletin , 31, 443–453. (2005) <http://dx.doi.org/10.1177/0146167204271651>

Long, K. and Spears, R. (1997) The self-esteem hypothesis revisited: Differentiation and the disaffected . In R. Spears, ed. , P.J. Oakes, ed. , N. Ellemers, ed. and S.A. Haslam (eds), The Social Psychology of Stereotyping and Group Life , pp. 296–317. Oxford: Blackwell.

Mael, F.A. and Tett, R.E. Identifying organizational identification . Educational and Psychological Measurement , 52, 813–824. (1992). <http://dx.doi.org/10.1177/0013164492052004002>

McGarty, C. Social identity theory does not maintain that identification produces bias and self-categorization theory does not maintain that salience is identification: Two comments on Mummendey, Klink and Brown . British Journal of Social Psychology , 40, 173–176. (2001) <http://dx.doi.org/10.1348/014466601164777>

Mullen, B., Brown, R.J. and Smith, C. Ingroup bias as a function of salience, relevance, and status: An integration . European Journal of Social Psychology , 22, 103–122. (1992) <http://dx.doi.org/10.1002/ejsp.2420220202>

Mummendey, A. and Schreiber, H.J. Better or just different? Positive social identity by discrimination against, or by differentiation from outgroups . European Journal of Social Psychology , 13, 389–397. (1983) <http://dx.doi.org/10.1002/ejsp.2420130406>

Mummendey, A. and Schreiber, H.J. 'Different' just means 'better': Some obvious and some hidden pathways to ingroup favoritism . *British Journal of Social Psychology* , 23, 363–368. (1984) <http://dx.doi.org/10.1111/j.2044-8309.1984.tb00652.x>

Oakes, P.J. (1987) The salience of social categories . In J.C. Turner, ed. , M.A. Hogg, ed. , P.J. Oakes, ed. , S.D. Reicher, ed. and M.S. Wetherell (eds), *Rediscovering the Social Group: A Self-categorization Theory* , pp. 117–141. Oxford: Blackwell.

Oakes, P.J., Haslam, S.A. and Turner, J.C. (1994) *Stereotyping and Social Reality* . Oxford: Blackwell.

Oakes, P.J. and Turner, J.C. Is limited information processing capacity the cause of social stereotyping? *European Review of Social Psychology* , 1, 111–135. (1990) <http://dx.doi.org/10.1080/14792779108401859>

Ouwerkerk, J.W., Ellemers, N. and De Gilder, D. (1999) Group commitment and individual effort in experimental and organizational contexts . In N. Ellemers, ed. , R. Spears, ed. and B.J. Doosje (eds), *Social Identity: Context, Commitment, Content* , pp. 184–204. Oxford: Blackwell.

Postmes, T. (2003) A social identity approach to communication in organizations . In S.A. Haslam, ed. , D. Van Knippenberg, ed. , M.J. Platow, ed. and N. Ellemers (eds), *Social Identity at Work: Developing Theory for Organizational Practice* , pp. 81–98. New York: Psychology Press.

Postmes, T. and Branscombe, N.R. (2010) *Rediscovering Social Identity: Core Sources* . New York: Psychology Press.

Postmes, T., Haslam, S.A. and Swaab, R. Social influence in small groups: An interactive model of identity formation . *European Review of Social Psychology* , 16, 1–42. (2005) <http://dx.doi.org/10.1080/10463280440000062>

Postmes, T., ed. and Jetten, J. (eds) (2006) *Individuality and the Group: Advances in Social Identity* . London: Sage. <http://dx.doi.org/10.4135/9781446211946>

Rabbie, J.M., Schot, J.C. and Visser, L. Social identity theory: A conceptual and empirical critique from the perspective of a behavioral interaction model . European Journal of Social Psychology , 19, 171–202. (1989) <http://dx.doi.org/10.1002/ejsp.2420190302>

Reicher, S.D. (1987) Crowd behaviour as social action . In J.C.Turner, M.A.Hogg, P.J.Oakes, S.D.Reicher and M.S.Wetherell, Rediscovering the Social Group: A Self-categorization Theory , pp. 171–202. Oxford: Blackwell.

Reicher, S.D. and Haslam, S.A. Rethinking the psychology of tyranny: The BBC Prison Study . British Journal of Social Psychology , 45, 1–40. (2006)

Reicher, S.D., Haslam, S.A. and Hopkins, N. Social identity and the dynamics of leadership: Leaders and followers as collaborative agents in the transformation of social reality . Leadership Quarterly , 16, 547–568. (2005) <http://dx.doi.org/10.1016/j.lequa.2005.06.007>

Reicher, S.D., Spears, R. and Haslam, S.A. (2010) The social identity approach in social psychology . In M.S. Wetherell, ed. and C.T. Mohanty (eds), The Sage Handbook of Identities . London: Sage. <http://dx.doi.org/10.4135/9781446200889>

Rink, F. and Ellemers, N. Diversity as a source of common identity: Towards a social identity framework for studying the effects of diversity in organizations . British Journal of Management , 18, S17-S27. (2007) <http://dx.doi.org/10.1111/j.1467-8551.2007.00523.x>

Rubin, M. and Hewstone, M. Social identity theory's self-esteem hypothesis: A review and some suggestions for clarification . Personality and Social Psychology Review , 2, 40–62. (1998) http://dx.doi.org/10.1207/s15327957pspr0201_3

Ryan, M.K. and Haslam, S.A. The Glass Cliff: Exploring the dynamics surrounding the appointment of women precarious leadership positions . Academy of Management Review , 32, 549–572. (2007) <http://dx.doi.org/10.5465/AMR.2007.24351856>

Sachdev, I. and Bourhis, R.Y. Power and status differentials in minority and majority group relations . European Journal of Social Psychology , 21, 1–24. (1991) <http://dx.doi.org/10.1002/ejsp.2420210102>

Scheepers, D. and Ellemers, N. When the pressure is up: The assessment of social identity threat in low and high status groups . Journal of Experimental Social Psychology , 41, 192–200. (2005) <http://dx.doi.org/10.1016/j.jesp.2004.06.002>

Schmitt, M.T. and Branscombe, N.R. The internal and external causal loci of attributions to prejudice . Personality and Social Psychology Bulletin , 28, 620–628. (2002) <http://dx.doi.org/10.1177/0146167202288006>

Sherif, M. (1967) Group Conflict and Co-operation: Their Social Psychology . London: Routledge and Kegan Paul.

Simon, B. The perception of ingroup and outgroup homogeneity: Re-introducing the intergroup context . European Review of Social Psychology , 3, 1–30. (1992) <http://dx.doi.org/10.1080/14792779243000005>

Simon, B. and Brown, R. Perceived intragroup homogeneity in minority-majority contexts . Journal of Personality and Social Psychology , 53, 703–711. (1987) <http://dx.doi.org/10.1037/0022-3514.53.4.703>

Simon, B. and Klandermans, B. Politicized collective identity: A social psychological analysis . American Psychologist , 56, 319–331. (2001) <http://dx.doi.org/10.1037/0003-066X.56.4.319>

Smith, H.J., Tyler, T.R. and Huo, Y.J. (2003) Interpersonal treatment, social identity, and organizational behavior . In S.A. Haslam, ed. , D. van Knippenberg, ed. , M. Platnow, ed. and N. Ellemers (eds), Social Identity at Work: Developing Theory for Organizational Practice , pp. 155–172. London: Psychology Press.

Spears, R. and Manstead, A.S.R. The social context of stereotyping and differentiation . European Journal of Social Psychology , 19, 101–121. (1989) <http://dx.doi.org/10.1002/ejsp.2420190203>

Tajfel, H. Cognitive aspects of prejudice . *Journal of Social Issues* , 25, 79–97. (1969) <http://dx.doi.org/10.1111/j.1540-4560.1969.tb00620.x>

Tajfel, H. Social identity and intergroup behaviour . *Social Science Information* , 13, 65–93. (1974) <http://dx.doi.org/10.1177/053901847401300204>

Tajfel, H. The exit of social mobility and the voice of social change . *Social Science Information* , 14, 101–118. (1975) <http://dx.doi.org/10.1177/053901847501400204>

Tajfel, H. (1978a) Interindividual behaviour and intergroup behaviour . In H. Tajfel (ed.), *Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations* , pp. 27–60. London: Academic Press.

Tajfel, H. (1978b) Social categorization, social identity, and social comparison . In H. Tajfel (ed.), *Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations* , pp. 61–76. London: Academic Press.

Tajfel, H. (1978c) The achievement of group differentiation . In H. Tajfel (ed.), *Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations* , pp. 77–98. London: Academic Press.

Tajfel, H. Social psychology of intergroup relations . *Annual Review of Psychology* , 33, 1–39. (1982) <http://dx.doi.org/10.1146/annurev.ps.33.020182.000245>

Tajfel, H. and Billig, M. Familiarity and categorization in intergroup behavior . *Journal of Experimental Social Psychology* , 10, 159–170. (1974) <http://dx.doi.org/10.1016/0022-1031%2874%2990064-X>

Tajfel, H., Billig, M.G., Bundy, R.F. and Flament, C. Social categorization and intergroup behaviour . *European Journal of Social Psychology* , 1, 149–177. (1971) <http://dx.doi.org/10.1002/ejsp.2420010202>

Tajfel, H. and Turner, J.C. (1979) An integrative theory of intergroup conflict . In W.G. Austin, ed. and S. Worcher (eds), *The Social Psychology of Intergroup Relations* , pp. 33–47. Monterey, CA: Brooks/Cole.

Tajfel, H. and Turner, J.C. (1986) The social identity theory of intergroup behavior . In W.G. Austin, ed. , and S. Worchel (eds), *The Social Psychology of Intergroup Relations* , pp. 7–24. Monterey, CA: Brooks/Cole.

Tajfel, H. and Wilkes, A.L. Classification and quantitative judgement . *British Journal of Psychology* , 54, 101–114. (1963) <http://dx.doi.org/10.1111/j.2044-8295.1963.tb00865.x>

Turner, J.C. (1982) Towards a cognitive redefinition of the social group . In H. Tajfel (ed.), *Social Identity and Intergroup Relations* , pp. 15–40. Cambridge: Cambridge University Press.

Turner, J.C. Some comments on the measurement of social orientations in the minimal group paradigm . *European Journal of Social Psychology* , 13, 351–367. (1983) <http://dx.doi.org/10.1002/ejsp.2420130403>

Turner, J.C. (1985) Social categorization and the self-concept: A social cognitive theory of group behaviour . In E.J. Lawler (ed.), *Advances in Group Processes* , 2, 77–122. Greenwich, CT: JAI Press.

Turner, J.C. (1991) *Social Influence* . Milton Keynes: Open University Press.

Turner, J.C. (1999) Some current issues in research on social identity and self-categorization theories . In N. Ellemers, ed. , R. Spears, ed. , and B. Doosje (eds), *Social Identity: Context, Commitment, Content* , pp. 6–34. Oxford: Blackwell.

Turner, J.C. and Brown, R.J. (1978) Social status, cognitive alternatives and intergroup relations . In H. Tajfel (ed.), *Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations* , pp. 201–234. London: Academic Press.

Turner, J.C. and Haslam, S.A. (2001) Social identity, organizations and leadership . In M.E. Turner (ed.), *Groups at Work: Advances in Theory and Research* , pp. 25–65. Hillsdale, NJ: Erlbaum.

Turner, J.C., Hogg, M.A., Oakes, P.J., Reicher, S.D. and Wetherell, M.S. (1987) *Rediscovering the Social Group: A Self-categorization Theory* . Oxford: Blackwell.

Turner, J.C. and Oakes, P.J. (1997) The socially structured mind . In C. McGarty, ed. and S.A. Haslam (eds), *The Message of Social Psychology: Perspectives on Mind in Society* , pp. 355–373. Malden MA: Blackwell.

Turner, J.C., Oakes, P.J., Haslam, S.A. and McGarty, C.A. Self and collective: Cognition and social context . *Personality and Social Psychology Bulletin* , 20, 454–463. (1994) <http://dx.doi.org/10.1177/0146167294205002>

Turner, J.C. and Onorato, R. (1999) Social identity, personality and the self-concept: A self-categorization perspective . In T.R. Tyler, ed. , R. Kramer, ed. and O. John (eds), *The Psychology of the Social Self* . Hillsdale, NJ: Erlbaum.

Turner, J.C. and Reynolds, K.J. (2001) The social identity perspective in intergroup relations: Theories, themes and controversies . In R.J. Brown, ed. and S. Gaertner (eds), *Blackwell Handbook of Social Psychology*, Vol. 4: Intergroup Processes , pp. 133–152. Oxford: Blackwell.

Turner, J.C., Reynolds, K.J., Haslam, S.A. and Veenstra, K. (2006) Reconceptualizing personality: Producing individuality through defining the personal self . In T. Postmes, ed. and J. Jetten (eds), *Individuality and the Group: Advances in Social Identity* , pp. 11–36. London: Sage. <http://dx.doi.org/10.4135/9781446211946>

Tyler, T.R. and Blader, S.L. (2000) Cooperation in Groups: Procedural Justice, Social Identity, and Behavioral Engagement . Philadelphia: Psychology Press.

Van Dick, R. (2004) My job is my castle: Identification in organizational contexts . In C.L. Cooper, ed. and I.T. Robertson (eds), *International Review of Industrial and Organizational Psychology* , 19, 171–204. Chichester: Wiley.

Wetherell, M.S. (1987) Social identity and group polarization . In J.C. Turner, ed. , M.A. Hogg, ed. , P.J. Oakes, ed. , S.D. Reicher, ed. and M.S. Wetherell (eds), *Rediscovering the Social Group: A Self-categorization Theory* , pp. 142–170. Oxford: Blackwell.

Wright, S.C. (2000) Strategic collective action: Social psychology and social change . In R. Brown, ed. and S. Gaertner (eds), *Blackwell Handbook of Social Psychology: Intergroup Processes* , pp. 223–256. Oxford: Blackwell.

<http://dx.doi.org/10.4135/9781446249222.n45>

Handbook of Theories of Social Psychology

Self-Categorization Theory

Contributors: John C. Turner & Katherine J. Reynolds

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Self-Categorization Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n46>

Print pages: 399-418

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n46>

[p. 399 ↓]

Chapter 46: Self-Categorization Theory

Abstract

The focus of this chapter is self-categorization theory (SCT). SCT is a theory of the nature of the self that recognizes that perceivers are both individuals and group member, explains how and when people will define themselves as individual and group entities and its implications, and examines the impact of this variability in self-perception ('I' to 'we') for understandings of mind and behaviour. As a result, it has generated a range of distinctive subtheories, hypotheses and findings across a range of significant areas in social psychology. This chapter outlines central steps in the theory's development, its unique contribution and the impact of its ideas with specific details provided in the areas of social influence (more recently, leadership and power) and individuality (e.g. personal self, personal self-perception, personal self-beliefs). In the final section, the way SCT can be applied to better understand and solve a range of social issues is highlighted. A specific example is provided of how core SCT ideas are being implemented in secondary schools with the aim of improving school outcomes (e.g. learning, bullying, wellbeing). It is our view that through an understanding of SCT (and related work) it is possible to appreciate the important and distinctive contribution of social psychology to other areas of psychology and cognate fields.

Introduction

This chapter is focused on self-categorization theory (SCT), its development, distinctive ideas, intellectual contribution and applicability to social issues. Given that the founder of SCT (the first author of this chapter) was a cofounder of social identity theory (SIT) with Henri Tajfel, there is much that these theoretical perspectives co-contribute to understanding and debates in social psychology. To appreciate what is distinctive about

SCT it is necessary to some degree to examine aspects of SIT (see also Haslam and Ellemers, [Chapter 45](#), this volume). As far as possible, though, this chapter will focus on SCT, acknowledging where relevant overlaps and common themes. This chapter can only provide an overview of core points. There are other more detailed accounts of the beginnings and contribution of SCT (e.g. Turner, 1987a, 1996; Turner and Oakes, 1989, 1997; Turner and Reynolds, 2010; Turner et al., 1987, 1994).

The proponents of both SIT and SCT are vocal in arguing that social psychology must [p. 400 ↓] acknowledge the functional interdependence of mind and society in its theorizing about the nature of mental processes (Turner and Oakes, 1997). People live, work and act in a socially structured system, where there are group-based regularities of perception, cognition and conduct and this reality has psychological consequences. SIT and SCT capture the socially embedded, situated, shared, social, group-located properties of human beings. This view contrasts with other approaches that reduce the working of the mental system to general (individual) psychological properties (e.g. information processing and memory systems) or the asocial (social environment-free) nature of the individual perceiver (e.g. personality, biology).

Building on the work of Lewin, Asch, Sherif and others it is argued that human beings are both individuals and group members, that they have personal and group aspects. Both theories argue that the psychological nature of individuals (e.g. the self, mind, cognition, information processing, memory, behaviour) has to be apprehended within an understanding of groups and membership in society. SIT and SCT define the proper and defining task of social psychology as studying and proposing theories consistent with the interplay between psychological functioning and the socially and/or culturally shared properties of human life (e.g. What does social life tell us about the mind? How does the mind make social interaction and society possible? How is the mind affected by social life?; see Turner and Oakes, 1997 for a more detailed discussion).

Theories in social psychology offer an approximation of reality that can be further investigated, elaborated and refined to provide a consistent explanation of the class of phenomena of interest. Effectiveness and parsimony typically are the dimensions on which theories are assessed. Thinking about SCT in this way the phenomena of interest is to understand, explain and predict how people come to think, feel and act as a psychological group and, importantly, the circumstances when this will occur

and its consequences. Through understanding the cognitive definition of the self, *how* and *when* perceivers define themselves and others as individual and group entities, SCT explains when a group is ‘a group’. The theory is at the centre of explaining the way the individual mind makes possible, and is impacted by, the fact that human beings are social animals (Turner and Oakes, 1997). SCT aims to be an effective and parsimonious theory of the self-process which contributes to explaining the functioning of the mind and behaviour.

There is a large body of work that has investigated the workings of the theory and derivations in an immense range of issues in the field (and beyond) including intergroup relations and prejudice, the nature of the group and the psychological basis of group and collective processes, social influence processes such as conformity, group polarization, minority influence, consensualization and leadership, crowd behaviour, social cooperation, group cohesion, social cognition (stereotyping, categorization), collective action and social change, the nature of the self, communication and language, and, latterly, the personal self, individuality and personality processes. In fact, many chapters in this volume engage with fundamental SCT concepts and ideas. It is also the case that implications of this theory extend beyond social psychology to psychology at large (and especially the problem of cognition) and the other social sciences (Haslam et al., 2010; Postmes and Branscombe, 2010; Reynolds et al., 2010). So how did this theory develop, why is it important and what is its impact on a range of social issues? We now turn to explore and explain the theory in more detail starting with the history of its development and then its core aspects and contribution.

SCT: Personal Narrative of its Development

In this section, the early beginnings of SCT are described. More formally, the development [p. 401 ↓] of the contemporary theory broadly can be summarized as involving three main steps (see also Turner and Reynolds, 2010). The first, was the distinction between personal identity and social identity and the hypothesis that it is social identity that is the basis of group behaviour. The second step, which occurred while Turner was at the Institute of Advanced Studies (IAS) at Princeton in 1982–1983, involved both an elaboration of the personal-social identity distinction to levels of self-categorization (e.g., individual, subgroup, superordinate), and the formalization of

the theory (Turner, 1985). The third step, conducted mainly in the 1980s and 1990s, involved a systematic program of research on the self-concept and stereotyping. What emerged was a more detailed and integrated understanding of the nature of the self and its implications for the foundation of cognition (Oakes et al., 1994; Turner and Oakes, 1997; Turner and Onorato, 1999; Turner et al., 1994).

Social Identity and the Psychology of the Group

The story of SCT begins in 1971 when John Turner started his PhD under Henri Tajfel's supervision at the University of Bristol in the UK. Like SIT, SCT begins with the minimal group studies published in that year, in the first volume of the *European Journal of Social Psychology*. The minimal group data had shown that social categorization into groups, in isolation from and unconfounded by all the variables normally thought to cause group formation and negative intergroup attitudes (interpersonal interdependence, history of conflict), was sufficient for discrimination. Individuals assigned more of a resource to others who were in the same group as themselves (ingroup) compared to members of a group which did not include them (outgroup). Furthermore, participants acted in ways that maximized the difference in allocations between the two groups even at the expense of allocating maximum resources to the group to which they belonged. SIT was concerned with explaining why subjects discriminated in the minimal group paradigm. SCT addressed a different question: Why did subjects identify with the minimal groups at all and act in ways that reflected that these group identities mattered to them?

On Turner's arrival at Bristol in September 1971, Tajfel's explanation of the minimal group findings had progressed from there being a 'generic norm' of ingroup favouritism or ethnocentrism (Tajfel, 1972). In a French textbook on social categorization, Tajfel had offered a new explanation of the findings. In this chapter, Tajfel introduces and defines the concept of social identity, outlines that groups provide their members with a positive social identity and that such positivity derives through establishing a valued distinctiveness for their own groups compared with other groups. Turner's first task was to flesh out this social identity and positive ingroup distinctiveness explanation

of the minimal group data. He reviewed the role of social categorization in intergroup relations and the findings of the minimal group paradigm, producing a review paper written before the end of 1971. This paper was presented by Turner at the Small Group Meeting of the European Association of Experimental Social Psychology (EAESP) on Intergroup Relations held at Bristol in February 1972. The paper, which was eventually published with additional data in 1975, showed how a systematic account of minimal and other forms of intergroup discrimination and ingroup bias (in terms of a process he called social competition), could be provided using social identity processes and not necessarily conflict of interests (e.g. Sherif, 1967).

The new analysis was summarized by Tajfel as the 'social categorization-social identity-social comparison-positive distinctiveness' sequence (Tajfel, 1974, 1978). The sequence provided a theoretical framework for understanding intergroup behaviour. Social categorizations defined people's place in society and through being internalized into the self, together with their emotional and value [p. 402 ↓] significance, provided people with social identities. Through social comparison on dimensions associated subjectively with perceivers' social values these social identities could be evaluated and provide valued positive distinctiveness for one's group (compared with other groups). The motive for positive distinctiveness could lead, under certain conditions, to ingroup favouring intergroup responses. At no time was it argued that ethnocentrism was universal or that social categorization automatically and inevitably produced ingroup bias or favouritism. If this were the case there would be no need for the development of theory to explain when such outcomes were more or less likely to define social relationships (Tajfel and Turner, 1979).

Examination of these processes was an important focus of work during the early to mid seventies at the University of Bristol. During this time, Tajfel proposed a continuum of human behaviour framed at one end by interpersonal behaviour and at the other by intergroup behaviour (Tajfel, 1978). Tajfel referred to the continuum as 'acting in terms of self' versus 'acting in terms of group'. The shift along the continuum is associated with distinct forms of social behaviour. At the interpersonal end, it is expected that there should be variability in behaviour towards ingroup and outgroup members. As the social situation nears the intergroup end, though, attitudes and behaviour become more grouplike or uniform.

The continuum was important because it highlighted that group behaviour and social identity were expected only under selected conditions and motivated more work to be done specifying the social psychological conditions that lead to group rather than individual attitudes and actions. Variables such as the permeability, legitimacy and the stability of status differences between groups in a particular social system were identified as shaping whether a situation would be characterized by consensual intergroup or interpersonal behaviour. The continuum also allowed Tajfel and Turner to make a distinction between acting as an individual and acting as a group member while at the same time recognizing that people were capable of both. Work on social categorization could be further developed to specify what this meant psychologically. It was this task that became the focus of SCT.

So while Tajfel and Turner continued their work (with others) on social identity, intergroup relations and social change, that culminated in a series of influential papers (e.g. Tajfel, 1978; Tajfel and Turner, 1979; Turner and Brown, 1978), Turner from 1978 onwards also focused on the psychological processes that underpin movement along the behavioural continuum. At a conference in 1978 at the University of Rennes in France held by the European Laboratory of Social Psychology (LEPS) he presented a paper entitled 'Towards a cognitive redefinition of the social group' which explained ideas on the psychological group (Turner, 1982). Turner developed a causal analysis of the psychological process related to movement along the interpersonal–intergroup continuum. He suggested that an individual's self-concept comprised definitions of self that included both personal identity and social identity. Social identity (self-definition in terms of social category memberships) was explicitly distinguished from personal identity (self-descriptions in terms of personal and idiosyncratic attributes) and situational variations in the self-concept were recognized with the idea that social identity could function at the relative exclusion of personal identity.

Turner proposed a theory of group behaviour in terms of an 'identity mechanism' to explain movement along the interpersonal–intergroup continuum. He hypothesized that as people defined themselves and others as members of the same category, they would self-stereotype in relation to the category and tend to see themselves as more alike in terms of the defining attributes of the category. This process is referred to as depersonalization. It was argued that it is 'the cognitive redefinition of the self – from unique attributes and individual differences to shared social category memberships

and associated [p. 403 ↓] stereotypes – that mediates group behaviour' (Turner, 1984: 528). It explains how individuals can psychologically be group members and 'reinstates the group as a psychological reality and not merely a convenient label for describing the outcome of interpersonal processes' (Turner, 1984: 535). This identity mechanism transforms the interpersonal–intergroup continuum into a cognitive, social psychological theory of the group (Turner, 1985, Turner et al., 1987; Turner et al., 1994).

Having applied for and received funding in 1978 for the new theory, Turner and his research group (Wetherell, Smith, Reicher, Oakes, Hogg, Colvin in roles as research assistants, PhD students or both) started applying these fundamental ideas in various areas. Initially the focus was social influence (conformity, group polarization, influence within the crowd), psychological group formation and the distinction between personal and group-based attraction (trying to show how group cohesion was a function of social identification rather than interpersonal attraction), and the problem of the salience of social categories.

Levels of Self-Categorization and Formalization of the Theory

While Turner was at IAS Princeton (1982–1983) he conceptualized further the categorization processes at work in personal identity and social identity. The focus was on the workings of self-categorization processes and the cognitive grouping of the self as being similar to some class of stimuli in contrast to some other class of stimuli. Ideas by Rosch (e.g. 1978) and others were particularly useful in thinking about processes of class inclusion and levels of inclusiveness. The personal–social identity distinction was reformulated as levels of self-categorization where people can define or categorize themselves at different levels of abstraction; for example, at the interpersonal level (where self is defined as a unique individual relative to others available for comparison), at the intergroup level (where self is defined as being a group member in contrast to a relevant outgroup), and at the superordinate level (where self is defined as a human being in contrast to other lifeforms). Self-categorizations at levels less inclusive than the individual person are also possible (e.g. intrapersonal comparisons within defining the personal self; for example, Reynolds and Turner, 2006). A central idea is that lower-

order self-categories were formed *inter alia* from social comparisons within higher-order ones and higher-order ones were formed *inter alia* on the basis of lower level ones (for more detail see Turner et al., 2006).

As part of the theory's development it was necessary to address the issue of what determines which identity emerges in a given situation (e.g., personal identity or social identity and the specific content of these). It was in Oakes' PhD on the salience of social categories, that Oakes and Turner addressed this issue (Oakes, 1987). Bruner's (1957) analysis of categorization and perception was adapted to correspond to the social domain. Bruner argued that 'all perceptual experience is necessarily the end product of a categorization process' (1957: 124). He held a functional view of categorization where the determinants of cognitive accessibility were a function of contextual factors and the current goals, needs and purposes of the perceiver. He used the formula of 'relative accessibility \times fit' to describe the conditions under which a stimulus was captured by a category and given meaning by the perceiver. The aim was to provide the perceiver with the information they needed to make sense of a stimulus and at the point when they needed to know it.

In this SCT work on salience, Oakes and Turner originally defined normative fit as the degree to which perceived similarities and differences between group members correlated with the social meaning of group memberships and in a direction consistent with such meaning of the group identities (e.g. it is expected that men and women differ in relation to independence and dependence and that the pattern of interaction in the given [p. 404 ↓] situation between men and women is consistent with men being independent and women being dependent; Oakes, 1987). Another aspect of salience was comparative fit. Defining comparative fit also was related to another project where Wetherell and Turner, in Wetherell's PhD, were developing a social identity explanation of group polarization (Wetherell, 1987). While at Princeton, Turner was trying to provide a quantitative principle that would allow some way of explaining why groups would polarize as a function of individuals' pretest views on any issue in any given context. The aim was to predict which person or position would become prototypical (representative, or most defining of the group) and when that prototype would polarize or not. Turner succeeded with the development of the metacontrast principle (Turner, 1985; Turner and Oakes, 1986, 1989).

The principle states that a collection of individuals tend to be categorized as a group to the degree *inter alia* that the perceived differences between them are less than the perceived differences between them and other people (outgroups) in the comparative context. As an example, in a given situation men will be categorized as independent and women as dependent when the differences between women and men in relation to this dimension are greater than those amongst the men and amongst the women available for comparison. Furthermore, any specific person or position tends to be seen as more prototypical of the group as a whole to the degree that the perceived differences between that person and other ingroup members are less than the perceived differences between that person and outgroup members.

Using principles of accessibility (based on Bruner) and fit (comparative and normative) it is possible to explain which of many identities will guide perception and behaviour in any given context. The central insight is that if the meaning given to a situation (including the self) is an outcome of categorization processes that are inherently comparative, then self-categories also are infinitely variable, contextual and relative.

Revisiting the Self-Concept and Stereotyping

A direct implication of the SCT analysis is that a self-category could not be stored as a fixed, cognitive structure in some mental system before it was used waiting to be activated (as Turner along with many others had thought originally). It became clear that basic understandings of the functioning of the cognitive system (e.g. memory, perception, information processing, stereotyping) and the self-concept (e.g. core and working self) had to be revisited.

The ideas at this time were facilitated by the work of Tajfel, Bruner and Rosch, but also Medin and Barsalou who argued that categories are expressions of theories and knowledge that explain how things go together ('meaning-making'; for example, Medin and Wattenmaker, 1987) and arguments against concepts as fixed mental models (e.g. Barsalou, 1987; see Oakes et al., 1994; Turner and Reynolds, 2010). Based on SCT it is argued that the variability of self-categories is central to how the perceiver (as an

individual and group member) responds in a world that also is variable and dynamic. Which group becomes salient for people, when, and its associated content or meaning, changes as a function of interactions between individuals and groups and the dynamic nature of such interactions. Shifts in self-categorization and the content of group-based judgments of oneself and others (e.g. stereotyping) reveals how self-categories are oriented to reality in which there are both individuals and groups in continuous dynamic interaction.

As an example of this point let us consider a person's stereotypes that men are independent and women are dependent. These stereotypes have to be understood within the broader intergroup relationship between men and women in society and shared understandings of that relationship. At times this intergroup relationship and social comparison as 'males' compared to 'females' will be particularly salient shaping people's social identity and attitudes and behaviours in a given situation. [p. 405 ↓] It is also the case that a fixed stereotype formed and stored as 'women are dependent' and 'men are independent' will not serve the perceiver well in the face of changes in the relationship between men and women in society. To be functional for the perceiver the cognitive process needs to be able to represent new and emerging understandings of intergroup relations and be responsive to social change processes. If this were not the case the cognitive system would be impoverished and not very adaptive to its environment.

Along these lines, research conducted at this time in the theory's evolution demonstrated that stereotypes are not rigid and erroneous but reflect perceptions of group relations from the perceiver's (possibly variable) vantage point. Likewise, one's self-concept (personal and collective) is flexible and responsive to contextual stability and variability. It became clear as expected that different self-categories can become salient (e.g., myself as an individual, woman or Australian) and the content of a particular category can change as a function of the salient comparative context (Australians compared to Americans/Australians compared to Chinese) and ongoing change (e.g., the historically evolving nature of what it means to be Australian).

To summarise this phase of theoretical development, then, it is argued that the self-category is a variable judgment formed on the basis of categorization-in-context. A person brings to a situation relatively enduring knowledge about the self (personal and

collective), and this information is used as a psychological resource in a given situation. This knowledge, in interaction with contextual factors, then produces a particular self-categorization and associated attitudes and behaviours. It is also the case that this knowledge (one's perceiver readiness) can be updated as a function of current self-categorizations and the accessibility of certain knowledge (and its meaning) can change as a result of the same processes.

To bring the points highlighted in the above narrative together it is possible to summarize the core theoretical developments in SCT as follows;

- 1. As with SIT it is argued that humans are not merely individuals and neither are our minds. Individuals, groups and intergroup relations exist. Human beings are both individuals and group members and therefore have both personal identity and social identity. Furthermore, based on SCT the psychological depersonalization of the self in terms of social identity produced 'group behaviour' and emergent group processes (e.g. influence, cooperation, cohesiveness). Conversely, defining oneself in terms of an idiosyncratic personal identity, in terms of individual differences from others and distinctive personal attributes, produces 'individual behaviour'.
- 2. People can define or categorize themselves at different levels of abstraction. It is possible to define oneself as an individual, as a member of particular groups in contrast to others and as a member of higher-order more inclusive groups. More inclusive self-categories define what is socially negotiated and affirmed as being valued, appropriate and right. At different times in different situations we define the self in different ways and such variation in the relative salience is seen as normal and ever-present.
- 3. Salience explains the way a particular situation (that includes the self) is categorized and given 'meaning'. The way the situation is categorized and understood by the perceiver will determine both self-perception and behaviour. Salience is a function of an interaction between the perceiver's readiness to use a self-category in a given instance and the fit of that self-category to the apprehended stimulus reality.

These three ideas in combination summarize core aspects of SCT (see Turner, 1987a). We now turn to outline the way these ideas are impacting in two specific areas: social

influence, which includes work on leadership and power; and individuality, which includes work on personal identity and personality processes.

SCT: Its Ideas and Intellectual History

Given the volume of work that speaks to the intellectual contribution of SCT to the [p. 406 ↓] field, in this section the focus will be on outlining in more detail two areas only: social influence and individuality. Social influence is an area that was a focus for initial work in SCT that has been extended to provide a new analysis of leadership and power (e.g. Turner, 1991; 2005; Turner and Haslam, 2001; Turner et al., 2008). Social influence itself is at the centre of social psychology with many significant theories in the field addressing the scientific study of how people come to influence one another affecting their attitudes, affect and actions. A more systematic consideration of personal identity, individuality and personality processes is an emerging area of inquiry (over the last 5 years or so) where the scope and relevance of SCT currently is being investigated (e.g. Reynolds and Turner, 2006; Reynolds et al., 2010; Turner et al., 2006). There is also a fundamental connection and interplay between these two areas. It is argued that it is through social identity processes and associated social influence (from others who are similar and 'like us') that group norms, values and beliefs can come to affect those individuals who define themselves in terms of those groups. This work, then, examines more closely the interplay between the group and the individual person.

SCT and Social Influence

To engage with the analysis of social influence offered by SCT, it is necessary to recognize the way the categorization process is understood within this theory and in particular the workings of the metacontrast principle. To reiterate, all things being equal, a collection of individuals (stimuli) tend to be categorized as a group (cognitively placed into the same class) to the degree that the perceived differences between them are less than the perceived differences between them and other people (outgroups) in the context of interest.

A number of studies have demonstrated how a psychological group emerges using these principles of categorization (see Haslam and Turner, 1992). Hogg and Turner (1987), for example, showed that when people were organized into mixed-sex groups (men and women) or same-sex groups (men-only or women-only), individuals were more likely to define themselves in terms of gender and to accentuate their similarity to those of the same gender in the mixed-sex settings as opposed to when only men or women were present (see also Oakes et al., 1991).

When people are considered to be in the same class of stimuli ('us' rather than 'them') they are cognitively grouped as similar perceivers confronting the same stimulus situation. This similarity leads people to tend to agree; it also creates an expectation that they ought to agree and respond in the same way (in reactions, judgement, attitudes, behaviour) and motivates people to bring about such agreement. In terms of explaining more specifically how 'others' come to affect one's own attitudes and behaviour, the stages are summarized by Turner (1987b) as follows:

- 1. Individuals define themselves as members of a distinct social category.
- 2. They learn or develop the appropriate, expected, desirable behaviours that are correlated with category membership, and differentiate it from other categories (e.g. the stereotypical norm).
- 3. They assign the norms and attributes of the category to themselves (internalization) through the process of depersonalization and self-stereotyping.
- 4. Their behaviour therefore becomes normative as their category membership becomes more salient.

Internalization is critical to the emergent social norms having an impact on one's attitudes and behaviour (see also Kelman, 1958, 2006) and is affected by the degree to which individuals consider themselves psychologically to be members of the particular group.

Haslam et al. (1999) provide one demonstration of aspects of these processes showing that it is precisely through social identity that idiosyncratic views become socially organized and consensual. Participants were asked to reflect on their social category membership [p. 407 ↓] as Australians (social identity condition) or to focus on their

uniqueness from others (personal identity condition) and completed a checklist identifying words typical of people from Australia before and after a group interaction phase. This manipulation of social identification did result in participants indicating that their national identity was more important to them in the social identity compared with the personal identity condition. There also was evidence that participants defined themselves in terms of the social category membership. In the social identity condition (and especially postinteraction), there was greater consensus on the attributes that defined the category of Australian, as well as the emergence of different content to describe the stereotypical attributes of the category.

Turner (1987b, 1991) also argued that subjective validity, certainty, competence, correctness and so on (e.g. what is considered factual and accurate), is a direct function of similar others in the same stimulus situation being understood to agree with one's own response. It is this point that transformed understandings of informational and normative influence into the one process of referent informational influence (Turner, 1991). Because other ingroup members are viewed as similar to oneself, they become a valid source of information and a testing ground for one's own views on relevant dimensions. Under these conditions, other group members can come to have an impact on one's own thoughts, attitudes and behaviours. It is this process of social influence that is important in explaining how others 'like us' play an important role in shaping the psychology of the person. Both certainty and uncertainty are related to the degree to which 'similar others' are perceived to agree or disagree with one's own response and are an outcome of the workings of the categorization process.

Turner (1987b, 1991) outlines a range of strategies to address situations where there is disagreement with others defined as being 'similar' including (a) changing our views in line with ingroup opinion, (b) attempts to influence other ingroup members to adopt a different stance through processes of mutual influence, (c) recategorization of ingroup members as being outgroup and (d) clarification of the stimulus situation (i.e. ensuring that reference is being made to the same thing; David and Turner, 1996, 1999; McGarty et al., 1994; Turner, 1991). It is argued that it is only within a shared ingroup framework that differences in perspective (e.g. criticism, new ideas, deviance) can be resolved through discussion, clarification and mutual influence. Through these processes ingroup members can shape each others' norms, values and beliefs in significant ways (re)defining 'who we are' and 'what we do'.

Much of the empirical work on these social influence processes has focused on showing not just that ingroup members are more influential than outgroup members but that the definition of who is included in the ingroup and who is excluded is a dynamic outcome of the workings of the categorization process (Haslam et al., 1992). One example of this point concerns shifts in levels of inclusiveness where hitherto subgroups are recategorized in relation to a higher-order superordinate ingroup along with associated empathy, trust, co-operation, positivity and all the other qualities that follow perceptions of self-other similarity and being 'ingroup'. Who is included in the 'ingroup' and who is excluded can be redefined shifting both the 'meaning' of the group (its defined content and norms) and who has opportunities for influence within the group. Extremists, for example, within a group can gain or lose influence as a function of the outgroup against which the ingroup defines itself (e.g. Haslam and Turner, 1995).

In relation to the social influence process, these ideas have been demonstrated, refined and documented, in particular, in the area of minority and majority influence. In the work of David and Turner (e.g. 1996, 1999), there has been a focus on the SCT principles underlying social influence and engagement with both majority and minority influence (see also Moscovici, 1976; Turner 1991). [p. 408 ↓] In one of the David and Turner studies (1996, Study 1), participants (either proconservationists or prologgers) indicated their attitudes to logging prior to an influence message and immediately after the influence attempt and three weeks after the attempt. Participants were presented with a prologging or proconservation message from the 'Friends of the Timber Industry' or 'Friends of the Forest', respectively. The message was presented as representing the majority or minority position within the timber industry or conservationists. In this way, participants received an influence attempt from an ingroup majority, ingroup minority, outgroup majority or outgroup minority.

The findings revealed that when the source of the message was outgroup irrespective of whether it was majority or minority, participants shifted away from the position advocated by the source – there was not social influence. In the ingroup conditions, participants moved in the direction of the source in both the majority and minority conditions immediately following the influence attempt. In line with Moscovici (1980) the condition that revealed the most long-term shift or change was when the message was attributed to an ingroup minority source. There is additional work showing that as a hitherto outgroup is recategorized as part of a more inclusive ingroup (as a function of

the frame of reference shaping the judgements of similarity and difference) it is possible for these members to exert greater influence. The implications of these results for the influence field more broadly are discussed in detail elsewhere (e.g. see David and Turner, 2001; Turner, 1991).

Additional empirical work has investigated the argument that categorization of similarities and differences between stimuli (people) not only leads to the formation of classes (ingroup and outgroup) but defines the relative prototype of the group. There is a hierarchy of relative influence that will follow the hierarchy of members' perceived relative prototypicality: where a specific person (or position) tends to be seen as more prototypical of the group when the perceived differences between the person and other ingroup members is less than the perceived differences between that person and outgroup members. There are direct links between the influence hierarchy and notions of leadership with respect to who will be influential in a group and be able to affect others to willingly engage in certain activities and behaviours. It follows that group members will emerge as leaders (those with the most influence) to the degree that they are perceived as relatively prototypical of the group as a whole (and in ways that fit existing normative expectations with respect to leadership) and that the most prototypical person will tend to be recognized as the leader where such a role is defined.

What has flowed from this analysis of social influence is a fundamentally new understanding of leadership and power (e.g. Haslam et al., 2011; Turner, 2005; Turner and Haslam, 2001; Turner et al., 2008). Leadership within SCT is conceptualized as a group process related to relative influence and power within a group (e.g. Turner and Haslam, 2001; Haslam, 2004). The breakthrough idea is that leadership rests on an individual's ability to be seen as prototypical of a shared social identity and hence will have greater influence as a result of such categorization processes. Influence over other group members becomes possible when leaders are seen as embodying 'who we are' (and in ways that normatively fit expectations of 'our' leaders).

In line with these points, power as the ability to have impact through others also rests on group identity and influence processes (Turner, 2005). It is through social identity processes that leaders are able to get others willingly to exert their will and as such mobilize 'followers' to action to achieve certain 'projects' (including the coercion of

those who are not on board). In this SCT analysis group identity and the associated willing support of followers it enables, allows groups to gain the resources they require to achieve their shared goals. These ideas are supported by a range of experimental and field studies [p. 409 ↓] which show that ingroup leaders (and those that are more versus less prototypical) have more potential to influence their followers, are perceived as more effective, are trusted more, and are seen as more charismatic (e.g. see Haslam, 2004; Suba#ic' et al., 2011; van Knippenberg et al., 2004).

People follow leaders because they embody ‘us’, and define what ‘we’ think is true and right, and do a better job than the rest of us of expressing what ‘we’ have in common and what we seek to achieve collectively. There also potentially are individual factors at play, but they exert influence only insofar as they are seen at any time by any given group as representing its identity better than others do. Some leaders are ‘identity entrepreneurs’ who through engaging in argumentation and political rhetoric seek to maintain their relative prototypicality and their position (e.g. Reicher and Hopkins, 2001). There is also evidence that leaders can attempt to restructure the social context or frame of reference and the definition of the group in ways that make their position more prototypical. Seeking conflict with an outgroup is one such response (e.g. Rabbie and Bekkers, 1978). The same is true when one demonizes, scapegoats and discriminates against a minority (sub)group. Prejudice against a minority can be used to reshape the mainstream identity, put one at the core, and increase one's influence (Turner, 2005; Turner et al., 2008).

Thus understanding leadership as a group process does not deny the capacity of certain leaders to make use of their insights into that process. The point is that leadership is an ability to genuinely influence and it is an outcome of group identity rather than being linked to the preordained life trajectory of any one individual (e.g. Haslam, 2004; Reicher et al., 2005; Turner and Haslam, 2001; Turner, 1991). It is through defining the group identity that leaders are able to position themselves in ways that maximize their influence and impact on ‘what we do’. In a more general sense, though, it should be apparent that it is through the construction of definitions of ‘who we are’ and ‘who we are not’ and associated social influence that people's opinion, norms, attitudes and behaviours can become consensualized, coordinated and transformed into collective action.

SCT and Individuality

SCT's theoretical analysis of the nature of the self and self-process also has implications for understanding personal identity, individuality and personality processes. The first point is that a key contribution of the SCT is that the social comparative features that define one's social identity in a given context can also be applied to understand one's self-definition as an individual (Haslam et al., 2010; Oakes et al., 1994; Reynolds and Turner, 2006; Turner et al., 2006). A critical idea is that whether impressions, perceptions and judgements of oneself and others are group-based or individuated, depends on the levels of abstraction at which the categorization process operates (which is a function of the goals and motives of the perceiver and the elements of the situation being cognized). Rather than personal identity reflecting the relatively stable and enduring features of an individual, the nature of individuality is forged through categorization and social comparison. This argument means that one's sense of who they are as an individual can vary depending on the social comparative context.

The point to emphasise is that one's values (beliefs, norms, worldviews) are variable and socially mediated and defined by ingroup memberships and relevant social influence processes. Under certain conditions, they also become a referent through which one's distinctiveness from others can be defined and emphasized. The content that is generated to describe personal identity depends on some comparative reference and this can result in different (or the same) self-descriptors being generated depending on the context. In a sense, individual differences can be thought about theoretically in this framework as *relative* individual differences because categorization and 'meaning' involves [p. 410 ↓] comparison and contrast (Onorato and Turner, 2004; Turner and Onorato, 1999).

Examining these arguments is complex because the aim is to develop a detailed theoretical analysis of the nature of the self-process (e.g. formation, functioning, consequences, (dis)continuity), but, people bring with them their already defined experiences as individual and group members and often function in relatively stable personal and group contexts. Exploring the mechanisms or processes that might explain self-stability and self-change and its consequences in contexts where for most people there is much stability in their group and personal experiences is a challenge.

There is much that can be done, though, to investigate these ideas empirically. A central theme of this work to date is to demonstrate the workings of the categorization process in relation to personal identity processes. Along these lines, Mavor, Reynolds and Skorich (2010) have investigated the impact of having people complete self-ratings in contexts where self and others are evaluated alone (intrapersonal context) or in comparison to each other (interpersonal context). At the group level there is evidence that one's own group is viewed as being more variable and heterogeneous when the group is judged alone (an intragroup context) rather than in comparison to a relevant outgroup (an intergroup context; for example, Haslam et al., 1995). It is also the case that personal self-judgments can vary depending on features of the comparative context (e.g. intrapersonal versus interpersonal). Thus if individuals compare themselves to others (interpersonal) rather than making assessments in isolation (intrapersonal), they are more likely to characterize themselves in a dispositional way. The interpersonal context accentuates the similarities and differences between the person and comparison other, leading to a strong sense of one's self-defining features. In this way, the comparative context has an impact on personal self-categorizations, and such categorizations also can be variable depending on the frame of reference (Guimond et al., 2007).

In addition, there is a growing body of evidence showing the impact social identity processes can have on a range of outcomes often associated more with individual-level characteristics and abilities (cognitive performance, wellbeing, self-reported personality). Work on social identity or stereotype threat shows that when one's social identity is salient and the stereotype of the group on the dimension of interest is negative, this can have an impact on cognitive ability (e.g. intelligence) and performance on dimensions relevant to the meaning or stereotype of the group (Steele and Aronson, 1995). Reicher and Haslam (2006) examined the impact of group processes and social identity on a range of more clinical outcomes (e.g. depression, anxiety, paranoia). Williams et al. (2008) have research findings that show that contamination anxiety (an aspect of obsessive-compulsive disorder) is affected not only by the ethnic social category of the respondent (e.g. African American or European American), but by whether the ethnic identity is salient or not when completing the anxiety measure. Bizumic et al. (2009) show that social identity is significantly related to, and mediated the relationship between, organizational factors and individual

psychological wellbeing (e.g. self-esteem, positive affect and job involvement, but also negative aspects such as depression, anxiety, loss of emotional control and aggressive and disruptive behaviour).

More specific investigations also are ongoing in relation to personality and people's self-reported sense of what characterizes and defines them as a person (e.g., self-beliefs, the Big Five). In personality theory and research, there is increasing recognition that one's social roles (e.g. daughter, worker) can impact on self-rated personality (Roberts and Donahue, 1994). The norms, expectations and meanings associated with certain roles can become internalized into the self-concept shaping a person's sense of self. An important element of this process is the impact of social interactions and the function of others' expectations, reactions and appraisals in shaping one's own beliefs about oneself (Roberts and Caspi, 2003).

[p. 411 ↓]

It is argued that the nexus between one's roles, identity and personality could be a force for continuity through, for example, the selection of environments that are consistent with and affirm one's self view (e.g. Swann and Read, 1981). It is also possible that personality may change through exposure to new roles that provide opportunities to engage in novel behaviours. The new roles could be associated with stages of normal adult development (e.g. parenthood, joining the workforce) or actively sought as people seek to improve, develop and reframe their personhood (e.g. more like their ideal image of themselves). Roberts and Mroczek (2008) argue that the findings that personality traits continue to change across the life course highlights the need for further work on the causes and mechanisms responsible for such change.

Based on the self-categorization theory of social identity and social influence, such life-development change (and broader social changes) may well affect one's group memberships and associated social identities. As different people come to be defined as similar to oneself, they offer new opportunities for social influence and the potential for one's theories, expectations and beliefs about oneself and the world can change. The general point is that these social identity changes may well impact on personhood in significant ways (Reicher and Haslam, 2006; Reynolds et al., 2005; Reynolds and Turner, 2006; Turner et al., 2006).

In one preliminary study related to personality processes, participants complete standard personality measures at one point in time (phase 1) and also again under conditions where their non-Aboriginal Australian versus Aboriginal Australian social identity was made salient (phase 2). Results demonstrated that across time (approximately 8 weeks) there is a high level of consistency in participant-reported Neuroticism. There also was evidence of a significant impact of the social identity manipulation and one's identification as a non-Aboriginal Australian in explaining personality assessed at phase 2. Findings suggested that it was the depression subscale of the Neuroticism measure (Goldberg IPIP-NEO) that was impacted most strongly as a result of non-Aboriginal identity (Reynolds et al., 2011). It was explained that in this condition, comparisons between non-Aboriginal and Aboriginal Australian may have oriented participants towards collective emotions and stereotypes that are related to what has been a negative intergroup comparison in Australia's history (see Branscombe and Doojse, 2004).

These kinds of studies are designed to investigate the SCT analysis of how self-views or self-beliefs are (re)formed and in ways the recognise social identity processes and group factors. There is evidence that categorization and social comparison affects personal identity and that social identity processes can have an impact on cognitive performance, personality and well-being in ways consistent with theory. Such findings (although preliminary) indicate that group processes may well play a role in (trans) forming personhood in particular ways. There also is more work to be done examining the role one's individuality plays in shaping the nature, functioning and success (or otherwise) of groups. All of these questions flow from the theoretical analysis of the nature of the self offered by SCT.

SCT: Its Applicability to Social Issues

As the above discussions highlight, SCT provides novel and important insights into aspects of psychological function that span intergroup relations to individual functioning. Core theoretical ideas, then, have been applied to a range of areas in psychology many of which can be readily related to current social problems and issues. More specific examples are in the areas of antiracism and prejudice reduction (e.g. Gaertner et al., 1989), the dynamics of social stability and social change (e.g. Spears et al., 2002;

Suba#i# et al., 2008; Turner and Reynolds, 2003; Wright et al., 1990), the [p. 412 ↓] relationship between attitudes, social norms and behaviour (e.g. Goldstein et al., 2008; Terry and Hogg, 1996), organizational (group) processes such as identification, leadership, negotiation and conflict management, and working effectively with diversity (e.g. Ashforth and Mael, 1989; Haslam, 2004; Haslam et al., 2003; Hogg and Terry, 2000; Rink and Ellemers, 2007) and health and wellbeing related outcomes (e.g. Bizumic et al., 2009; Branscombe et al., 1999; Haslam et al., 2009). There is detailed work in these and other areas that outlines the specific contribution of SCT and the implications of the approach. In the space available, one more recent project will be outlined in detail to give a flavour of the way SCT theoretical ideas (and related work) are being used both to understand and define certain social problems and implement novel solution.

Currently, social psychologists at the Australian National University are involved in a joint project with the local Department of Education concerned with applying core SCT ideas to improving school outcomes such as numeracy and literacy, attendance, challenging behaviour and staff and student wellbeing (Bizumic et al., 2009; Reynolds et al., 2007). Based on the arguments outlined above, as people come to define themselves as group members they should be more willing to internalize the norms and values of the group, act in line with these norms and be influenced by those that are most representative of the group. The aim of the project is to affect core aspects of individual functioning (learning, wellbeing, bullying/aggression) through making changes to the norms, values and beliefs that define the school as a whole (superordinate level) and relationships between groups (conflictual or cooperative) within the school environment (subgroup level). It is argued that to the degree these 'interventions' affect one's psychological connection to the school (school identification) and understandings of what it means to be school members (social identity content) there should be an impact on school outcome measures.

There are a number of strategies that can be implemented to affect social identity processes and to make higher-order identities more salient and thereby unify members in a common purpose and affect intergroup relations within the school setting. It is possible, for example, to (a) clarify the school's (organization's) shared mission and in essence what differentiates the school from others (i.e. what makes us 'us', what are 'our' goals), (b) restructure the way the school functions creating new structures that

shape which groups and divisions are likely to become meaningful psychologically (e.g., activities structured by year group are likely to affect the salience of group memberships defined by year groups), and (c) increase the extent to which members participate and are involved in decisions that affect them, which in turn affects their identification with the group, 'ownership' of decisions and willingness (intrinsically) to enact them (Tyler and Blader, 2000).

Building on these points, an initial starting point in applying SCT in schools has been to build a sense of shared mission. Staff, students and interested parents and community members (as subgroups) have been involved in a process where the vision, purpose and ideal behaviours for staff and students within a particular school have been identified (e.g. Haslam et al., 2003). The collated information has been endorsed by the relevant parties and communicated to clarify the norms, values and beliefs that define the school (at school assemblies, in the classroom, on posters displayed around the school). A whole range of school activities and functions are shaped by this sense of 'who we are' (e.g. professional development, codifying shared practices, celebration of achievements, championing individuals who exemplify the school's mission).

In some of the schools, the aim has been to better integrate the school values with the school structure so as to promote more positive cooperative relationships between the subgroups within the school (e.g. junior and senior school, staff and students). At one school, in order to reduce the division amongst [p. 413 ↓] staff (across faculties) and amongst staff and students and between year groups of students, a pastoral house care system was introduced, in which other categorizations crossed through being 'house' members (Crisp et al., 2001). Effectively such efforts serve to reduce the fit between certain group memberships and certain attributes and ways of functioning and introduce the possibility of other meaningful identities emerging to shape behaviour (e.g. staff do not just interact within their faculty but also across school planning).

In another school the focus has been on the classroom culture and shifting relationships, from one in which the teacher relies on coercion and extrinsic motivation to manage relationships with students and achieve learning outcomes, to one focused on leadership, influence and building intrinsic motivation (Turner, 2005). It is argued that the ability of one individual (a teacher) to get another party (a student or group of students) to willingly engage in some task or activity is a leadership process. Learning

requires, at least in part, a process of social influence to emerge between the teacher and students. In order to achieve this, in the classroom the teacher is encouraged to seek to involve students in decision making about their learning and to reach shared consensus on learning goals and standards. The class is also involved in deciding on the process through which they all will achieve certain learning outcomes. As a result, it is more likely for students to 'own', feel responsible for, and be intrinsically motivated to achieve, certain outcomes and also be more likely to support each other in achieving what is now a shared collective enterprise. Many of these ideas are consistent with initiatives in the educational context (including the quality learning movement) but locating these ideas within a broader theory of psychological functioning provides a more integrated approach and serves to reinforce the importance of certain educational initiatives over others.

The impact of initiatives and interventions such as these are being assessed on a range of school outcomes using a longitudinal design across a time period of up to 4 years. Although the SCT-based interventions are in the early stages of being introduced, initial results are in line with predictions. There is evidence that social (school) identification is significantly related to, and mediates the relationship between, organizational factors and individual psychological wellbeing (Bizumic et al., 2009). Organizational factors include the degree to which staff and student support the goals and objectives of the school, endorsement of school leadership and decision-making processes, the academic emphasis within the school and the fairness and clarity of rules and consistency in their implementation. These factors often form aspects of school climate measures in the educational domain. Measures of wellbeing address positive aspects of personal functioning, such as self-esteem, positive affect and job involvement, but also negative aspects, such as depression, anxiety, loss of emotional control and aggressive and disruptive behaviour (e.g. bullying, attention seeking, victimization, spreading rumours, social exclusion). The covariation of these measures suggests that if changes are made to schools which boost one's sense of psychological connection or belonging to the group, wellbeing and challenging behaviour should also be affected.

This work and the preliminary findings are exciting for a number of reasons. First, they highlight the relevance of social psychology in addressing issues in both clinical and educational contexts (e.g. wellbeing, aggression/bullying in schools). Second, the findings reinforce the need to integrate further the role of social identity processes in

understanding the (individual) psychology of the person. Third, the work speaks to the importance of recognizing all aspects of human psychological functioning (personal and social) in addressing social issues and problems. It is argued that there is added value in the definition of issues and the development of solutions that recognize that people are both individuals and group members and target [p. 414 ↓] the most appropriate level in relation to the issue at hand.

Conclusion

In this chapter, core aspects of SCT have been outlined. This theory is part of a history of ideas in social psychology where there is a rejection that the person and their psychology is bound up with 'basic processes' that somehow sit apart from social experience, interaction and group life. The challenge has been to develop a model of human psychological functioning that engages with the group and society to show both how being social has affected the workings of the human mind (e.g., thoughts, emotions, memory, perception, imagination) and how the workings of the human mind make the social possible. Through a detailed analysis of the basic processes that underlie the psychological group and the cognitive definition of the self, SCT offers a non-reductionist view of the mind which has generated a range of distinctive subtheories, hypotheses and findings across a range of significant areas in social psychology. In this way the theory has demonstrated both its effectiveness and parsimony. This task has not been easy; it has been one that has involved the efforts of many and it is one that is not yet fully completed (Turner and Reynolds, 2010). It is our view that through serious engagement with the nature of the self and self-categorization process as defined in SCT it will be possible to advance social psychology and understanding of human psychology.

Acknowledgements

This chapter was supported by funding from the Australian Research Council to both authors, including an Australian Research Fellowship to Dr Reynolds and an Australian

Professional Fellowship to Professor Turner. We would like to thank Paul van Lange for his very helpful comments on the draft manuscript.

John C.Turner Katherine J.Reynolds

References

Ashforth, B.E. and Mael, F. Social identity theory and the organization . Academy of Management Review , 14, 20–39. (1989)

Barsalou, L.W. (1987) The instability of graded structure: Implications for the nature of concepts . In U. Neisser (ed.), Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization , pp. 101–140. Cambridge: Cambridge University Press.

Bizumic, B., Reynolds, K.J., Turner, J.C., Bromhead, D. and Suba#i#, E. The role of the group in individual functioning: School identification and the psychological well-being of staff and students . Applied Psychology: An International Review , 58, 171–192. (2009) <http://dx.doi.org/10.1111/j.1464-0597.2008.00387.x>

Branscombe, N.R. and Doosje, B. (2004) Collective Guilt: International Perspectives . New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139106931>

Branscombe, N.R., Schmitt, M.T. and Harvey, R.D. Perceiving pervasive discrimination among African-Americans: Implications for group identification and well-being . Journal of Personality and Social Psychology , 77, 135–149. (1999) <http://dx.doi.org/10.1037/0022-3514.77.1.135>

Bruner, J.S. (1957) Going beyond the information given . In J.S. Bruner, ed. , E. Brunswik, ed. , L. Festinger, ed. , F. Heider, ed. , K.F. Muenzinger, ed. , C.E. Osgood, ed. , and D. Rapaport (eds), Contemporary Approaches to Cognition , pp. 41–69. Cambridge, MA: Harvard University Press. (Reprinted in Bruner, J.S. [1973]. Beyond the Information Given pp. 218–238. New York: Norton).

Crisp, R.J., Hewstone, M. and Rubin, M. Does multiple categorization reduce intergroup bias? *Personality and Social Psychology Bulletin* , 27, 76–89. (2001) <http://dx.doi.org/10.1177/0146167201271007>

David, B. and Turner, J.C. Studies in self-categorization and minority conversion: Is being a member of the outgroup an advantage? *British Journal of Social Psychology* , 35, 1–21. (1996) <http://dx.doi.org/10.1111/j.2044-8309.1996.tb01091.x>

David, B. and Turner, J.C. Studies in self-categorization and minority conversion: The ingroup minority in intergroup and intragroup social contexts . *British Journal of Social Psychology* , 38, 115–134. (1999) <http://dx.doi.org/10.1348/01446699164086>

David, B. and Turner, J.C. (2001) Self-categorization principles underlying majority and minority influence . In J.P. Forgas, ed. and K.D. Williams (eds), *Social Influence: Direct and Indirect Processes* , pp. 293–313. Philadelphia: Psychology Press.

Gaertner, S.L., Mann, J., Murrell, A. and Dovidio, J.F. Reducing intergroup bias: The benefits of recategorization . *Journal of Personality and Social Psychology* , 57, 239–249. (1989) <http://dx.doi.org/10.1037/0022-3514.57.2.239>

Goldstein, N.J., Cialdini, R.B. and Griskevicius, V. A room with a viewpoint: Using social norms to motivate environmental conservation in hotels . *Journal of Consumer Research* , 35, 472–482. (2008) <http://dx.doi.org/10.1086/586910>

Guimond, S., Branscombe, N.R., Brunot, S., Buunk, B.P., Chatard, A., Désert, M., Garcia, D.M., Haque, S., Martinot, D. and Yzerbyt, V. Culture, gender, and the self: Variations and impact of social comparison processes . *Journal of Personality and Social Psychology* , 92, 1118–1134. (2007) <http://dx.doi.org/10.1037/0022-3514.92.6.1118>

Haslam, S.A. (2004) *Psychology in Organizations: The Social Identity Approach* , 2nd Edition. Thousand Oaks, CA: Sage Publications.

Haslam, S.A., Eggins, R.A. and Reynolds, K.J. The ASPIRe model: Actualizing Social and Personal Identity Resources to enhance organizational outcomes .

Journal of Occupational and Organizational Psychology , 76, 83–113. (2003) <http://dx.doi.org/10.1348/096317903321208907>

Haslam, S.A., Ellemers, N., Reicher, S., Reynolds, K.J. and Schmitt, M. (2010) The social identity perspective tomorrow . In T. Postmes, ed. and N. Branscombe (eds), Rediscovering Social Identity: Core Sources . New York: Psychology Press.

Haslam, S.A., Jetten, J., Postmes, T. and Haslam, C. Social identity, health and well-being: An emerging agenda for applied psychology . Applied Psychology: An International Review , 58, 1–23. (2009) <http://dx.doi.org/10.1111/j.1464-0597.2008.00379.x>

Haslam, S.A., Oakes, P.J., Reynolds, K.J. and Turner, J.C. Social identity salience and the emergence of stereotype consensus . Personality and Social Psychology Bulletin , 25, 809–818. (1999) <http://dx.doi.org/10.1177/0146167299025007004>

Haslam, S.A., Oakes, P.J., Turner, J.C. and McGarty, C. Social categorization and group homogeneity: changes in the perceived applicability of stereotype content as a function of comparative context and trait favourableness . British Journal of Social Psychology , 34, 139–160. (1995) <http://dx.doi.org/10.1111/j.2044-8309.1995.tb01054.x>

Haslam, S.A., Ellemers, N., Reicher, S., Reynolds, K.J. and Schmitt, M. (2010) Social identity tomorrow: Opportunities and avenues for advance . In T. Postmes, ed. and N. Branscombe (eds). Rediscovering Social Identity: Core Sources . pp. 357–379. Psychology Press.

Haslam, S.A., Reicher, S.D. and Platow, M.J. (2011) The New Psychology of Leadership: Identity, Influence and Power . London: Psychology Press.

Haslam, S.A. and Turner, J.C. Context-dependent variation in social stereotyping 2: The relationship between frame of reference, self-categorization and accentuation . European Journal of Social Psychology , 22, 251–277. (1992) <http://dx.doi.org/10.1002/ejsp.2420220305>

Haslam, S.A. and Turner, J.C. Context-dependent variation in social stereotyping 3: Extremism as a self-categorical basis for polarized judgement . European Journal of Social Psychology , 25, 341–371. (1995) <http://dx.doi.org/10.1002/ejsp.2420250307>

Haslam, S.A., Turner, J.C., Oakes, P.J., McGarty, C. and Hayes, B.K. Context-dependent variation in social stereotyping 1: The effects of intergroup relations as mediated by social change and frame of reference European Journal of Social Psychology , 22, 3–20. (1992) <http://dx.doi.org/10.1002/ejsp.2420220104>

Hogg, M.A. and Terry, D.J. Social identity and self-categorization processes in organizational contexts Academy of Management Review , 25, 121–140. (2000)

Hogg, M.A. and Turner, J.C. Intergroup behavior, self-stereotyping and the salience of social categories . British Journal of Social Psychology , 26, 325–340. (1987) <http://dx.doi.org/10.1111/j.2044-8309.1987.tb00795.x>

Mavor, K.I., Reynolds, K.J. and Skorich, D. (2010) Attributions for one's own behaviour . The role of comparison in explaining dispositional attributions and the actor-observer effect. Unpublished manuscript, The Australian National University.

McGarty, C., Haslam, S.A., Hutchinson, K.J. and Turner, J.C. The effects of salient group memberships on persuasion . Small Group Research , 25, 267–293. (1994) <http://dx.doi.org/10.1177/1046496494252007>

Medin, D.L. and Wattenmaker, W.D. (1987) Category cohesiveness, theories and cognitive archaeology . In U. Neisser (ed.), Concepts and Conceptual Development, Ecological and Intellectual Factors in Categorization , pp. 25–62. Cambridge: Cambridge University Press.

Mischel, W. Towards an integrative science of the person . Annual Review of Psychology , 55, 1–22. (2004) <http://dx.doi.org/10.1146/annurev.psych.55.042902.130709>

Moscovici, S. (1976) Social Influence and Social Change . London: Academic Press.

Moscovici, S. (1980) Towards a theory of conversion behaviour . In L. Berkowitz (ed.), Advances in Experimental Social Psychology , 13, 209–239. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960133-1>

Oakes, P.J. (1987) The salience of social categories . In J.C.Turner, M.A.Hogg, P.J.Oakes, S.D.Rieche, and M.S.Wetherell, Rediscovering the Social Group: A Self-Categorization Theory . pp. 117–141. Oxford: Blackwell.

Oakes, P.J., Haslam, S.A. and Turner, J.C. (1994) Stereotyping and Social Reality . Oxford: Blackwell.

Oakes, P.J., Turner, J.C. and Haslam, S.A. Perceiving people as group members: The role of fit in the salience of social categorizations . British Journal of Social Psychology , 30, 125–144. (1991) <http://dx.doi.org/10.1111/j.2044-8309.1991.tb00930.x>

Onorato, R.S. and Turner, J.C. Fluidity in the self-concept: The shift from personal to social identity . European Journal of Social Psychology , 34, 257–278. (2004) <http://dx.doi.org/10.1002/ejsp.195>

Postmes, T. and Branscombe, N. (2010) Sources of social identity . In T. Postmes, ed. and N. Branscombe (eds), Rediscovering Social Identity: Core Sources . pp. 1–13. Psychology Press.

Rabbie, J.M. and Bekkers, F. Threatened leadership and inter-group competition . European Journal of Social Psychology , 8, 9–20. (1978) <http://dx.doi.org/10.1002/ejsp.2420080103>

Reicher, S.D. and Haslam, S.A. Rethinking the psychology of tyranny: The BBC Prison Study . British Journal of Social Psychology , 45, 1–40. (2006)

Reicher, S.D., Haslam, S.A. and Hopkins, N. Social identity and the dynamics of leadership: Leaders and followers as collaborative agents in the transformation of social reality . Leadership Quarterly , 16, 547–568. (2005) <http://dx.doi.org/10.1016/j.lequa.2005.06.007>

Reicher, S. and Hopkins, N. (2001) *Self and Nation* . London: Sage. <http://dx.doi.org/10.4135/9781446220429>

Reynolds, K.J., Bizumic, B., Suba#i#, E., Melsom, K. and MacGregor, F. (2007) Understanding the school as an intergroup system: Implications for school reform and improving student and staff outcomes . Grant funded by the Australian Research Council and ACT Department of Education and Training . Canberra: The Australian National University.

Reynolds, K.J., Suba#i#, E., Bizumic, B., Turner, J.C., Branscombe, N. and Mavor, K.I. (2009) Social identity and personality processes: Non-aboriginal Australian identity and neuroticism . Manuscript in submission.

Reynolds, K.J. and Turner, J.C. Individuality and the prejudiced personality . European Review of Social Psychology , 17, 233–270. (2006) <http://dx.doi.org/10.1080/10463280601050880>

Reynolds, K.J., Turner, J.C., Branscombe, N. and Mavor, K.I. (2005) Self-categorization and personal identity: Integrating group and personality processes . Grant funded by the Australian Research Council . Canberra: The Australian National University.

Reynolds, K.J., Turner, J.C., Branscombe, N.R., Mavor, K.I., Bizumic, B. and Suba#i#, E. Interactionism in personality and social psychology: An integrated approach to understanding the mind and behaviour . European Journal of Personality , 24, 458–482. (2010) <http://dx.doi.org/10.1002/per.782>

Rink, F. and Ellemers, N. Diversity as a source of common identity: Towards a social identity framework for studying the effects of diversity in organizations . British Journal of Management , 18, 19–29. (2007) <http://dx.doi.org/10.1111/j.1467-8551.2007.00523.x>

Roberts, B.W. and Caspi, A. (2003) The cumulative continuity model of personality development: Striking a balance between continuity and change . In U. Staudinger, ed. and U. Lindenberger (eds), *Understanding Human Development: Life Span Psychology in Exchange with Other Disciplines* , pp. 183–214. Dordrecht: Kluwer. http://dx.doi.org/10.1007/978-1-4615-0357-6_9

Roberts, B.W. and Donahue, E.M. One personality, multiple selves: Integrating personality and social roles . *Journal of Personality* , 62, 201–218. (1994) <http://dx.doi.org/10.1111/j.1467-6494.1994.tb00291.x>

Roberts, B.W. and Mroczek, D.K. Personality trait stability and change . *Current Directions in Psychological Science* , 17, 31–35. (2008) <http://dx.doi.org/10.1016/j.jrp.2004.03.001>

Rosch, E. (1978) Principles of categorization . In E. Rosch, ed. and B.B. Lloyd (eds), *Cognition and Categorization* . Hillsdale, NJ: Erlbaum.

Sherif, M. (1967) Social Interaction Process and Products: Selected Essays . Chicago: Aldine Publishing.

Spears, R., Jetten, J. and Doosje, B. (2002) The (il) legitimacy of ingroup bias . In J.T. Jost, ed. and B. Major (eds), *The Psychology of Legitimacy* , pp. 332–362. UK and US: Cambridge University Press.

Steele, C.M. and Aronson, J. Stereotype threat and the intellectual test performance of African-Americans . *Journal of Personality and Social Psychology* , 69, 797–811. (1995) <http://dx.doi.org/10.1037/0022-3514.69.5.797>

Suba#i#, E., Reynolds, K. J. and Turner, J. C. The political solidarity model of social change: Dynamics of self-categorization in intergroup power relations . *Personality and Social Psychology Review* , 12, 330–352. (2008)

Suba#i#, E., Reynolds, K.J., Turner, J.C., Veentra, K. and Haslam, S.A. (2011) Leadership, power and the use of surveillance: Implications of shared social identity for leaders' capacity to influence . *Leadership Quarterly* .

Swann, W.B., Jr. and Read, S.J. Self-verification processes: How we sustain our self-conceptions . *Journal of Experimental Social Psychology* , 17, 351–372. (1981) <http://dx.doi.org/10.1016/0022-1031%2881%2990043-3>

Tajfel, H. (1972) Social categorization . In S. Moscovici (ed.), *Introduction à la psychologie sociale*, Vol. 1 . Paris: Larouse.

Tajfel, H. Social identity and intergroup behaviour . *Social Science Information* , 13, 66–93. (1974) <http://dx.doi.org/10.1177/053901847401300204>

Tajfel, H. (ed.) (1978) *Differentiation Between Social Groups: Studies in the Social Psychology of Inter group Relations* . London: Academic Press.

Tajfel, H. and Turner, J.C. (1979) An integrative theory of intergroup conflict . In W.G. Austin, ed. and S. Worchel (eds), *The Social Psychology of Intergroup Relations* , pp. 33–47. Monterey, CA: Brooks/Cole.

Terry, D.J. and Hogg, M.A. Group norms and the attitude-behavior relationship: A role for group identification . *Personality and Social Psychology Bulletin* , 22, 776–244. (1996) <http://dx.doi.org/10.1177/0146167296228002>

Turner, J.C. Social comparison and social identity: Some prospects for intergroup behaviour . *European Journal of Social Psychology* , 5, 5–34. (1975) <http://dx.doi.org/10.1002/ejsp.2420050102>

Turner, J.C. (1978) Social categorization and social discrimination in the minimal group paradigm . In H. Tajfel (ed.), *Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations* , pp. 27–60. London: Academic Press.

Turner, J.C. (1982) Towards a cognitive redefinition of the social group . In H. Tajfel (ed.), *Social Identity and Intergroup Relations* , pp. 15–40. Cambridge: Cambridge University Press and Paris: Editions de la Maison des Sciences de l'Homme.

Turner, J.C. (1984) Social identification and psychological group formation . In H. Tajfel (ed.), *The Social Dimension: European Developments in Social Psychology* , pp. 518–538. Cambridge: Cambridge University Press.

Turner, J.C. (1985) Social categorization and the self-concept: a social cognitive theory of group behaviour . In E.J. Lawler (ed.), *Advances in Group Processes* , pp. 77–122. Greenwich, CT: JAI Press.

Turner, J.C. (1987a) Introducing the problem: individual and group . *Rediscovering the Social Group: A Self-categorization Theory* , pp. 1–18. Oxford: Blackwell.

Turner, J.C. (1987b) The analysis of social influence . In J.C. Turner, ed. , M.A. Hogg, ed. , P.J. Oakes, ed. , S.D. Reicher, ed. and M.S. Wetherell (eds), Rediscovering the Social Group: A Self-categorization Theory , pp. 68–88. Oxford: Blackwell.

Turner, J.C. (1991) Social Influence . Milton Keynes, England: Open University Press and Pacific Grove, CA.: Brooks/Cole.

Turner, J.C. (1996) Henri Tajfel: An introduction . In W.P. Robinson (ed.), Social Groups and Identity: Developing the Legacy of Henri Tajfel , pp. 1–24. Oxford: Butterworth Heinemann.

Turner, J.C. Explaining the nature of power: A three-process theory . European Journal of Social Psychology , 35, 1–22. (2005) <http://dx.doi.org/10.1002/ejsp.244>

Turner, J.C. and Brown, R.J. (1978) Social status, cognitive alternatives and intergroup relations . In H. Tajfel (ed.), Differentiation Between Social Groups , pp. 201–234. London: Academic Press.

Turner, J.C. and Haslam, S.A. (2001) Social identity, organizations and leadership . In M.E. Turner (ed.), Groups at Work: Theory and Research , pp. 25–65. New Jersey: Lawrence Erlbaum Associates.

Turner, J.C., Hogg, M.A., Oakes, P.J., Reicher, S.D. and Wetherell, M.S. (1987) Rediscovering the Social Group: A Self-categorization Theory . Oxford and New York: Basil Blackwell.

Turner, J.C. and Oakes, P.J. The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence . British Journal of Social Psychology , 25, 237–252. (1986) <http://dx.doi.org/10.1111/j.2044-8309.1986.tb00732.x>

Turner, J.C. and Oakes, P.J. (1989) Self-categorization theory and social influence . In P.B. Paulus (ed.), The Psychology of Group Influence , pp. 233–275. Hillsdale, NJ: Erlbaum.

Turner, J.C. and Oakes, P.J. (1997) The socially structured mind . In C. McGarty, ed. and S.A. Haslam (eds), *The Message of Social Psychology* , pp. 355–373. Oxford: Blackwell.

Turner, J.C., Oakes, P.J., Haslam, S.A. and McGarty, C. Self and collective: Cognition and social context . *Personality and Social Psychology Bulletin* , 20, 454–463. (1994) <http://dx.doi.org/10.1177/0146167294205002>

Turner, J.C. and Onorato, R. (1999) Social identity, personality and the self-concept: A self-categorization perspective . In T.R. Tyler, ed. , R. Kramer, ed. and O. John (eds), *The Psychology of the Social Self* , pp. 11–46. Mahwah, NJ: Lawrence Erlbaum Associates.

Turner, J.C. and Reynolds, K.J. Why social dominance theory has been falsified . *British Journal of Social Psychology* , 42, 199–206. (2003) <http://dx.doi.org/10.1348/014466603322127184>

Turner, J.C. and Reynolds, K.J. (2010) The story of social identity . In T. Postmes, ed. and N. Branscombe (eds), *Rediscovering Social Identity: Core Sources* , pp. 13–32. Psychology Press.

Turner, J.C., Reynolds, K.J., Haslam, S.A. and Veenstra, K. (2006) Reconceptualizing personality: Producing individuality by defining the personal self . In T. Postmes, ed. and J. Jetten (eds), *Individuality and the Group: Advances in Social Identity* , pp. 11–36. London: Sage Publications. <http://dx.doi.org/10.4135/9781446211946>

Turner, J.C., Reynolds, K.J. and Suba#i#, E. (2008) Identity confers power: The new view of leadership in social psychology . In P. 't Hart, ed. and J. Uhr (eds), *Public Leadership: Perspectives and Practices* , pp. 52–72. Canberra: ANU E-press.

Tyler, T.R. and Blader, S. (2000) Cooperation in groups: Procedural justice, social identity, and behavioral engagement . Philadelphia, PA: Psychology Press.

van Knippenberg, D., van Knippenberg, B., De Cremer, D. and Hogg, M.A. Leadership, self and identity: A review and research agenda . *The Leadership Quarterly* , 15, 825–856. (2004) <http://dx.doi.org/10.1016/j.lequa.2004.09.002>

Wetherell, M. (1987) Group polarization . In J.C.Turner, M.A.Hogg, P.J.Oakes, S.D.Riecher and M.S.Wetherell, Rediscovering the Social Group: A Self-categorization Theory , pp. 142–170. Oxford: Blackwell.

Williams, M.T., Turkheimer, E., Magee, E. and Guterbock, T. The effects of race and racial priming on self-report of contamination anxiety . Personality and Individual Differences , 44, 744–755. (2008) <http://dx.doi.org/10.1016/j.paid.2007.10.009>

Wright, S.C., Taylor, D.M. and Moghaddam, F.M. Responding to membership in a disadvantaged group: From acceptance to collective protest . Journal of Personality and Social Psychology , 58, 994–1003. (1990) <http://dx.doi.org/10.1037/0022-3514.58.6.994>

<http://dx.doi.org/10.4135/9781446249222.n46>

Handbook of Theories of Social Psychology

Social Dominance Theory

Contributors: Jim Sidanius & Felicia Pratto

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Social Dominance Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n47>

Print pages: 418-439

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n47>

[p. 418 ↓]

Chapter 47: Social Dominance Theory

Abstract

This chapter outlines the intellectual and personal influences on the development of social dominance theory (SDT). SDT examines how societies organize themselves as group-based social hierarchies. SDT assumes that processes at different but intersecting levels of social organization, from prejudice to cultural legitimizing ideologies, produce and maintain hierarchical societal structure. The chapter examines the counteracting roles of hierarchy-enhancing and hierarchy-attenuating legitimizing ideologies and social institutions, the intersection between gender and arbitrary set discrimination (i.e., discrimination based on socially constructed group distinctions), the distinction between authoritarianism and social dominance orientation, and emphasizes the critical role of social power (as opposed to social status), and the need to see social dominance as an integrated and dynamic social system.

A Brief Overview of Social Dominance Theory

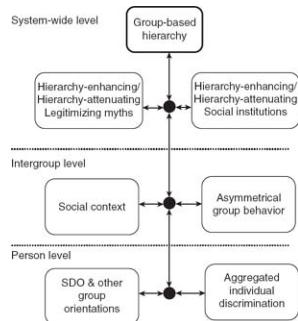
Stated most simply, social dominance theory (SDT) argues that intergroup oppression, discrimination, and prejudice are the means by which human societies organize themselves as group-based hierarchies, in which members of dominant groups secure a disproportionate share of the good things in life (e.g., powerful roles, good housing, good health), and members of subordinate groups receive a disproportionate share of the bad things in life (e.g., relatively poor housing and poor health). While the severity of group-based inequality varies across different societies and within any given society across time, the fact of group-based social hierarchy appears to be a human universal (e.g., Lenski, 1984). Because SDT attempts to describe the systematic processes that

form the dynamic system of societal inequality, its analysis considers the intersection of processes at multiple levels of social organization (see Pratto et al., 2006; Sidanius and Pratto, 1999 for recent reviews).

In a slight modification of Pierre van den Berghe's (1978) taxonomy of social categories, SDT observes that human group-based social hierarchies consist of three distinctly different stratification systems: (1) an *age-system*, in which adults and middle-age people have disproportionate social power over children and younger adults; (2) a *gender or patriarchal system* in which men have [p. 419 ↓] disproportionate social and political power compared to women; and (3) an *arbitrary-set system* in which socially constructed categories are hierarchically arranged. These arbitrary sets may be constructed to associate power and legitimacy with social categories like "race," caste, ethnicity, nationality, social class, religion, or any other group distinction that human interaction is capable of constructing. As the double-headed arrows in [Figure 47.1](#) are meant to indicate, we argue that group-based hierarchy both affects and is effected by roughly seven processes at three levels of analysis.

At the societal level the degree of group-based social hierarchy is effected by and affects two mutually antagonistic sets of forces: (1) hierarchy-enhancing and hierarchy-attenuating legitimizing ideologies, and (2) hierarchy-enhancing and hierarchy-attenuating social institutions. Hierarchy-enhancing and hierarchy-attenuating ideologies justify the establishment and maintenance of group-based social inequality or its exact opposite, respectively. To the degree that the relative balance of these opposing ideologies remains stable, the degree of social inequality remains stable over time, everything else being equal. Actions of hierarchy-enhancing and hierarchy-attenuating institutions also produce the level of inequality at the societal level. Hierarchy-enhancing social institutions allocate social resources to the advantage of dominant groups and to the disadvantage of subordinate groups, whereas hierarchy-attenuating social institutions have the opposite effect. Examples of hierarchy-enhancing institutions are internal security forces, large segments of the criminal justice system, and most large corporations. Examples of hierarchy-attenuating institutions are human rights and civil rights organizations, charities, and legal aid groups for the poor and the indigent (e.g., Sidanius et al., 1996).

Figure 47.1 An overview of social dominance theory



At the intergroup level, we posit two general processes that sustain inequality. First, aspects of unequal intergroup contexts afford prejudicial and discriminatory behavior. Unequal contexts readily dredge up stereotypes and remembered histories of past conflicts, perceived intergroup threat, and belief in separate identities, all of which provoke discrimination and stereotyping (see Pratto, 1999, for a review). Second, members of subordinate groups tend to behave in ways that are less beneficial to themselves and their ingroups than dominant group members do with reference to their ingroups. We call this *behavioral asymmetry*, and it is instantiated in many ways. For example, people in dominant groups follow their doctors' orders and study more than people in subordinate groups (see Sidanius and Pratto, 1999: 227–262). Behavioral asymmetry implies that group-based hierarchies are not solely maintained by the oppressive actions of dominants, but also by agency, albeit constrained agency, on the part of subordinates.

At the person level, the roles, prejudices, social beliefs that contribute to discrimination are coordinated, often in the same directions, so that thousands of aggregated individual acts of cruelty, oppression, and discrimination help sustain group-based hierarchy. Certain values, personality variables, [p. 420 ↓] political ideologies, and temperaments, including openness, conservatism, authoritarianism, and empathy make certain people more or less likely to be prejudiced or to discriminate against subordinates (Akrami and Ekehammar, 2006; Altemeyer, 1998; Pratto et al., 1994; Stephan and Finlay, 1999). In general, an individual's likelihood of performing hierarchy-enhancing or hierarchy-attenuating acts depends on her general desire to support and maintain group-based inequality, a characteristic we call social dominance orientation (SDO).

Thus, at core, there are three basic assumptions underpinning SDT. First, we assume that human social systems are dynamically tenacious. Thus, even as they adapt and change, societies that are group-based dominance hierarchies will tend to continually reorganize themselves, and even other societies, as such. Second, various forms of group-based oppression (e.g., sexism, racism, nationalism, ethnocentrism, classism) should be seen as specific instantiations of group-based social hierarchies. Third, the degree of group-based social hierarchy within any society at any given time will be the net result of the interaction of multileveled hierarchy-enhancing and hierarchy-attenuating forces within that society at any given time. Thus, the ultimate goal of SDT is to understand the multileveled processes which are responsible for the production, maintenance, and reproduction of group-based social hierarchy.

SDT: Early Roots and Personal Journeys

Jim's Personal Narrative

The basic building blocks of SDT were being assembled in my mind since childhood. As a 10-year-old boy growing up in New York City in the mid 1950s I had already become uncomfortably aware that being a "Negro" in America was not an altogether good thing. However, the deadly seriousness of this predicament did not become clear to me until the day I came across a *Jet* magazine article about a young Black man accused of whistling at a White woman in the South. The article described how a group of White men kidnapped this young Black man, castrated him, and poured gasoline onto the open gash where his genitals used to be. This story left a deep impression on me and I reread it over and over again, trying to grasp the meaning of such brutality.

Perhaps my most consciousness-altering confrontation with American racism occurred when I was a 16-year-old high school student. On my way home from school with my Jewish girlfriend and a White male friend, my male friend and I were followed into a public restroom in Highbridge Park by a White policeman, his gun drawn and demanding that we raise our hands and face the restroom wall. Not being aware of having broken any law, I asked the officer why we were being stopped. I was told

to shut the fu*k up, and marched off to the 33rd police precinct. Upon arrival I was told to sit down and once again, to “shut the fu*k up!” After some time had passed, I again demanded to know why I was being detained. This resulted in the arresting officer punching me in the face and yelling a string of racial epithets at me. I lost my composure and struck back. Immediately some four or five baton-wielding police officers pounced on me, beat me into near unconsciousness, placed me in overly tight handcuffs, and threw me into a holding cell. I then spent the night in a jail cell at the Brooklyn House of Detention. The next afternoon I was arraigned in criminal court and listened as the arresting officer testified that I was guilty of drunken disorderliness, interfering with traffic, and resisting arrest. Three witnesses disputed these claims (i.e., my girlfriend, my White male friend, and an independent witness to my arrest). After listening to all their testimony, the judge said that he would be lenient with me *this one time*. He then ordered my release, with the admonition that in future, I “show more respect for the law!”

[p. 421 ↓]

Even though I was released from custody, the message communicated to me was crystal clear. I was arrested, beaten, jailed, and arraigned for multiple acts of insubordination: insubordination for the crime of having a White girlfriend, insubordination for the act of questioning the legitimacy of my arrest, and most critically, insubordination by defending myself against physical attack by the police. By being told to “show more respect for the law,” I was clearly being told to keep my place, or else. This critical event led to a visceral understanding of the role the police and other armed authorities play in maintaining generalized submission and acquiescence from Black people in the US. Although this was the last time I was personally subject to police violence, I witnessed this kind of violence across many societies. The direct and vicarious experience of police violence influenced the development of SDT many years later.

Rather than reinforcing my submission to the American racial order, this early experience with the police had the opposite effect. I was transformed from a rather milk-toast liberal into an angry and resentful Black radical. After participating in numerous demonstrations and acts of resistance throughout the 1960s, I had finally had enough of American racism, and left the country in 1970, planning never to return. After traveling

to Canada, France, Germany, Denmark, and spending a few months in Algeria hanging out with some members of the Black Panther Party, I made my way to Sweden, where I eventually settled, raised a family, and was awarded a doctoral degree in political psychology.

The early years in Sweden were a revelation. While Swedes treated me with a certain degree of curiosity (at the time many Swedes had never seen a Black person in the flesh), their reactions to me were not laced with that combination of fear and loathing that had become such an intolerable part of my everyday experience with Whites in America.

Although my American origins very often shielded me from various slights and outright discrimination, it soon became clear that a number of other ethnic minorities were serving as targets of discrimination and devaluation (e.g., Finns, Turks, Roma). And so it was within every society I visited or learned anything about. These discrimination targets varied from people of sub-Saharan descent in Algeria, to Arabs from the Maghreb in France, to Turks in Germany and Denmark, to blond haired, blue-eyed Finns in Sweden, and to Roma in every country in Western and Eastern Europe. I also noticed an unsettling similarity in the manner in which the police treated members of these ethnic outgroups across the countries I visited. This treatment varied from a snarling intimidation to outright physical brutality, so reminiscent of my experiences with American police. Not only did I observe a thought-provoking cross-cultural similarity in the nature of police behavior towards ethnic minorities, the content of the stereotypes concerning these groups was also remarkably similar. Across a variety of different societies, the local ethnic subordinates were often described as lazy, conniving, criminal, dangerous, incompetent, and welfare-dependent.

In doing doctoral research in political psychology at the University of Stockholm, I came across a surprising and consistent finding which was to have a major influence on the later development of SDT. Namely, using two large and independent samples of Swedish high school students, my colleague, Bo Ekehammar, and I discovered some noteworthy differences in the sociopolitical attitudes of boys and girls, the strongest of which were the substantially higher levels of xenophobia and racism among boys than girls (see Ekehammar and Sidanius, 1982; Sidanius and Ekehammar, 1980, 1983). These findings were surprising because gender egalitarianism had been a major

component of Swedish political culture for half a century. While there was reason to expect attitudinal differences with respect to gendered issues (e.g., abortion rights), there was little reason to expect gender differences with respect to dimensions such as xenophobia [p. 422 ↓] and racism. Further, the higher levels of prejudice among boys were not moderated by differences in political ideology (Ekehammar, 1985); they were essentially of the same magnitude among communists as among fascists. Shortly after these findings were published, independent researchers replicated these results in other countries such as Great Britain and South Africa (see Furnham, 1985; Marjoribanks, 1981).

Because of the limited academic opportunities in Sweden, I decided to return to the US in 1983. I was initially comforted by the fact that the America I returned to was substantially less overtly racist than the America I had left behind 13 years earlier. However, it did not take too long for me to realize that beneath the surface of this increased racial inclusiveness, one could still clearly recognize a largely unchanged racial order underlying the bulk of social interactions. Despite the substantial progress achieved by the civil rights movement, it also became clear to me that this movement had failed in its central mission. The hierarchical racial order of American life remained very much as I had left it. The attempt to understand this sameness ensconced within change provided the initial emotional energy for the development of SDT.

Reading the history of reformist and revolutionary social movements, as well as the work of the neoclassical elitism scholars (e.g., Mosca, Pareto, Michels), convinced me that the failure of truly transformational change is the rule rather than the exception. Every attempt to replace group-based hierarchy with truly egalitarian social interaction has failed, without exception. These failures range from attempts at large-scale revolutionary transformation (e.g., the French, Russian, Mexican, Chinese, Vietnamese, Cambodian, and Cuban revolutions and the attempt at introducing economic democracy in Sweden during the mid 1970s), to small-scale attempts at egalitarian communities (e.g., the Oneida, Shaker, Harmonist, and Janssonist Communities of North America). While many of these revolutionary efforts have succeeded in replacing one group of ruling elites with another, and sometimes even decreasing the overall level of oppression, none have ever succeeded in their original goals of replacing group-based hierarchy with genuine egalitarianism.

While the building blocks of SDT lay scattered across disparate areas of my consciousness by the time I accepted a tenured position as Associate Professor of Psychology at UCLA in 1988, the first and rather underdeveloped form of SDT did not find its way onto paper until the summer of that year. Professor David O. Sears, one of my senior colleagues-to-be at UCLA, gave me a copy of his paper on symbolic racism he was to present at the upcoming meetings of the International Society of Political Psychology, and invited me to present a paper at this panel. I took the opportunity to react to David Sears' symbolic racism thesis. Rather than regard symbolic racism (defined as a combination of anti-Black affect and traditional American values such as self-reliance) as the ultimate source of White opposition to redistributive social policy favoring Blacks (e.g., busing, affirmative action), I argued that symbolic racism is better seen as one among several legitimizing ideologies serving the purpose of justifying the continued domination of Blacks by Whites, and more generally as the attempt of a dominant group to use a legitimizing ideology to maintain supremacy over a subordinate group. My rather incoherent and tedious reaction to Sears' paper was the primitive beginning of what was to grow into SDT (for a more coherent version of this initial argument, see Sidanius et al., 1992). However, the full development of SDT did not take place until I started to have theoretical jam sessions with Marilyn Brewer, a senior colleague, and distinguished intergroup relations specialist, and Felicia Pratto, a brilliant young woman I had first met when she was an undergraduate at Carnegie Mellon University, and with whom I later reconnected when she was a newly minted PhD from the social psychology program at New York University in 1989. [p. 423 ↓] Numerous critical conversations with Marilyn Brewer stimulated me to develop the central idea of the counterbalancing effects of hierarchy-enhancing versus hierarchy-attenuating social forces, while the collaboration with Felicia Pratto led to the conceptualization and initial measurement of the SDO construct, the conceptual and empirical distinctions between arbitrary-set and gender hierarchies, the extension of the person–environment fit perspective onto the psychology of intergroup relations, and several other aspects of SDT as it stands today.

Felicia's Personal Narrative

My story is not as dramatic as Jim's, but may give a sense of how a person can develop a consciousness of intergroup power, social exclusion, discrimination, legitimizing myths, and social justice.

My family was formed in the foothills of the Rockies by the intermarriages of second-generation Eastern and Southern European immigrants. Although almost no one acknowledges this in the presumption that all Whites are the same in the US, the early twentieth-century immigration produced a great deal of ethnic diversity and multilingualism. My family's story includes the fact that ethnic divides help stabilize power structures. My grandfather, Pete Pratto, born in 1900, became a coal miner for Colorado Fuel and Iron after he gave up being a cowboy and homesteader at age 40. Coal miners from many European countries, Mexico, and Japan lived in company towns, were paid poorly in company scrip, and were worked hard in dangerous conditions. The mining companies housed different ethnic groups in separate areas and did their best to stir up enmity as a way of preventing unionization from taking hold. Several of my grandfather's union activist friends were murdered, and this kind of intimidation and ethnic conflict prevented unions from taking hold for decades longer than they should have (e.g., Beshoar, 1957). My paternal grandmother, Bertha Bon, was a child at the Ludlow Massacre. Her family, and those of all the other miners on strike, was living in tents in the foothills because they were not allowed to live in company housing while on strike. On Orthodox Easter, 1914, their tents were firebombed and they were shot at by the Colorado state militia, which was working at the beck and call of John D. Rockefeller's Standard Oil. My grandmother told me that she hid behind a kitchen table turned on its side, and the only reason she felt she and her family survived is that a coal train went between where the militia shooters were and the families were cowering, allowing them to escape farther into the hills. With this family history, I could not grow up assuming that poor people are less virtuous than the rich, that hard work inevitably pays, nor that power is usually used for good and justice is always delivered.

In 1969 my family moved from the American West to Greensboro, NC, where my father was offered an academic job as a sociologist. This move from the West, where

everyone we knew was a “guy” because everyone had the same social class and was capable and friendly, taught me much about cultural ideologies as not only scripts but masks. We were told to expect “Southern hospitality,” but in fact the White neighbor kids yelled “damned Yankees” at my sister Anita and me as we walked home from school. Not much more subtly, my public school fourth-grade teacher, Mrs. Lambert, regularly asked my class to raise their hands and keep them raised if we were Jewish, then if we were Catholic, while she made notes that we never saw. Molly Ivins had a similar awakening about Southern culture when she was asked how a native Texan could grow up to be a progressive. When she had asked her mother about why she couldn’t drink out of the “colored” water fountain and her mother said that it was filthy, when Molly could see it was clean, Molly said she realized that if they were lying about race, they were probably lying about everything else too. So I did [p. 424 ↓] not feel welcome in North Carolina, but this led me to form great friendships with other “outsiders,” like African-Americans, immigrants, and Yankees. Thus, not everybody feels they belong, or is made to feel they belong, everywhere. This basic idea later found a place in SDT.

Attending college in Pittsburgh was refreshing. I enjoyed the mixture of working class and upper class kids from many states in a student culture where everyone could be creative and successful. Carnegie Mellon had its own foibles, but unlike the South, they weren’t designed to prevent certain people from serving in certain roles or to keep people on top who were not deserving, except for one problem: sexism.

In 1979 when I started, the student body was 70 percent men, and the science, engineering, and architecture schools had not only higher proportions, but also a very boyish culture. Dr. Goldberg, my wonderful physics section teacher who taught kids from the South and inner cities and got us to pass Physics 1, was the only PhD who taught a section (the others were led by graduate students). I do not know her whole story, but I do know that when I was a physics major, the non PhD instructor gave a full-letter-grade lower grade to every woman in the mechanics lab than he gave to her lab partner. We had thought about insisting on egalitarian relations with men in our personal lives, but not in institutions like schools. Given our instructor’s grading and his remarks about “girls” throughout the course, all but one of us women in physics chose to change majors or universities.

I had the fortune to earn my work-study money doing research for several social scientists at Carnegie Mellon, from whom I learned much. Susan T. Fiske, in particular, spent lots of time and effort mentoring me, entertaining my questions, teaching me how to do social psychology experiments, trying to teach my rough sensibilities on how to be a professional. After I graduated, she hired me to help on her research on stereotyping, which enabled me to support myself, grow up and save more money before graduate school, feel I belonged at Carnegie Mellon, and engage in peace activism. I was deeply vexed by social injustice, and so I thought I should study stereotyping because it was the only domain of social psychology interested in this problem. One afternoon I asked Susan why there were such inequities, such as the fact that I, with a BS, made exactly half as much as my husband-to-be did who had not finished his degree. That was “the market,” not really an explanation of why different kinds of work are valued differently. I asked why we have different people doing different kinds of jobs and Susan gave me the then-standard line about cognitive heuristics leading to stereotyping leading to discrimination. But this “how” answer to a “why” question was not satisfactory, and I blurted out, “What about” – searching for the missing concept – “POWER?” I suppose that is when I started to realize that even if scientists mainly answer how and not why questions, they at least should have complete descriptions of the processes. Still, because I was looking to elders about what to research, I did not follow through on my own intuitions until a few years later.

In graduate school at New York University (NYU), I had a ball doing experiments on automatic processing and stereotyping with John Bargh, and the social cognition process-focused atmosphere was very stimulating. But again, some questions remained unanswered. I recall that in some seminars, all the outcome measures correlated with participants' social class, and yet the papers were never trying to explain that relation. Also, during my third year, Trish Devine submitted her dissertation research for publication and made the same stir at NYU as elsewhere. She argued that stereotyping could be traced to more essential, unconscious, and uncontrollable cognitive processes (Devine, 1989). This work was an important lesson to me. This approach was in the air in social psychology – I saw Mahzarin Banaji's students at conferences and they contended that if we could only get at the automatic [p. 425 ↓] processes, we could eradicate discrimination (see Blair and Banaji, 1996, as an example of this work). Earlier, Hamilton and Gifford (1976) had shown that negative stereotypes about minority

groups could arise simply because of a nonmotivated, nonsocial cognitive process of associating infrequent features (group members and negative behaviors). This work is certainly intellectually elegant and the nonobviousness of the analysis was widely appreciated, including by me. But my read of social psychology at this time was that in erasing the *intention* from stereotyping and discrimination, the discipline also erased the fact that inequality has real *consequences*. Even outside the social cognition domain, our field had shifted from considering the consequences of racism for Black people (Clark and Clark, 1947), to whether focusing on racism would upset White people's self-image (Gaertner and Dovidio, 1986; Katz and Hass, 1988). So the study of racism had moved from being about how justice and equality could be realized by and for Black people, to how White people could be prevented from feeling uncomfortable. At a moment that I could have gone the cognitive–essentialist route in addressing social inequality intellectually, I decided this was just not the right approach. What I felt was essential to a theory of social inequality was (1) culture: the systems of meaning and ideologies that pattern behavior and social structure, (2) an overt acknowledgement of power as part of the social context, (3) a focus on consequential outcomes like inequality, and (4) a theoretical analysis that overtly showed how processes at different levels of analysis, from in the person, to intergroup, to society-wide to intersocietal, scaled. For such reasons I was very excited when Jim Sidanius showed me the rough sketch he had made of SDT just after I finished my PhD and was heading to California to join my new (and only) husband. My social cognition training served my work on SDT well because it has systematic methods to substantiate processes and link them together (this is what I tried to do in my pretenure work on SDT; Pratto, 1999). Through different personal and intellectual routes, Jim and I had come to similar sensibilities about what a real theory was and what had to be included in a theory of inequality.

Major Components of and Intellectual Influences on SDT

SDT has been influenced by a wide variety of perspectives both inside and outside of social psychology. These influences are all the more varied because of our different

training in personality, political history, and social cognition. The most important of these influences have come from:

- 1. authoritarian personality theory (a psychoanalytic approach to understanding the intersection between child-rearing practices, personality development, political ideology and prejudice; see Adorno et al., 1950);
- 2. early social identity theory (a psychological theory of intergroup discrimination composed of three basic elements: (a) social categorization, (b) psychological identification, and (c) social comparison and, if possible, the achievement of a positive comparison between an ingroup and an outgroup; see Tajfel and Turner, 1986);
- 3. Rokeach's two-value theory of political behavior (the notion that political behavior is a joint function of the value one places on both equality and freedom; see Rokeach, 1973);
- 4. Blumer's (1960) group position theory (the notion that racial prejudice is a result of attempts to establish and maintain favorable positions within a social hierarchy);
- 5. Marxism (Gramsci, 1971; Marx and Engels, 1846);
- 6. neoclassical elitism theory (or the notion that social hierarchies are ubiquitous and essentially inevitable; see Michels, 1911; Mosca, 1896; Pareto, 1901);
- 7. industrial/organizational psychology (Bretz and Judge, 1994); and
- 8. sociological work on institutional discrimination (Hood and Cordovil, 1992), cultural ideologies (e.g., Sanday, 1981), evolutionary biology (Trivers, 1972), evolutionary psychology (Betzig, 1993; van den Berghe, 1978) and biological anthropology (Dickemann, 1979).

[p. 426 ↓]

The Influence of Neo-Classical Elitism Theories and the Concept of Legitimizing Myths

One important idea that we borrowed from classical and neoclassical elitism theories concerns the nature of societal structure. With the decided exception of Marxism, these theories presume that social systems and complex social organizations are inherently hierarchically and oligarchically structured. Thus, what ordinarily passes for democratic rule on the surface is, in actuality, the exercise of control by economic and social elites (e.g., Dahl, 1989). For such reasons, we look not at a society's system of government, but rather its degree of group inequality and the mechanisms responsible for that inequality (e.g., Sidanius and Pratto, 1993).

The second major idea that many of these theories share in common concerns the role of ideas in producing and maintaining group-based inequality. Pareto (1935) argued that there are two major means by which members of dominant groups establish and maintain hegemony, force, and fraud. By force, Pareto simply meant the use or threat of physical force and intimidation. By fraud he referred to the use of consensually shared social ideology functioning to legitimize the dominant position of the powerful over the powerless. Elitism theories and Marxism acknowledge that physical intimidation is an important means by which dominants exploit and control subordinates, but they maintain that, in the long run, it is not the most effective means of social control. A more potent means of sustaining hierarchy is by controlling social legitimacy. Marxists refer to these legitimacy instruments as "ideology" and "false consciousness," Mosca refers them by the term "political formula," Pareto uses the notion of "derivations," and Gramsci invokes the idea of "ideological hegemony." All these assert that elites maintain control over subordinates by controlling what is and what is not considered legitimate discourse, and promoting the idea that the rule of elites is moral, just, necessary, inevitable, and fair. SDT calls these ideological instruments "legitimizing myths."

SDT defines legitimizing myths (LMs) as consensually shared ideologies (including stereotypes, attributions, cosmologies, predominant values or discourses, shared

representations, etc.) that organize and justify social relationships. LMs suggest how people and institutions should behave, why things are how they are, and how social value should be distributed. Because they are consensual and closely associated with the structure of their societies, LMs often have the appearance of being true. Consequently, those who reject them take risks and have work to do in explaining how and why they disagree.

Unlike Marxist approaches, including system justification theory (see Jost and Banaji, 1994), SDT does not assume that all such myths are false, nor that they only reinforce social hierarchy (Sidanius and Pratto, 1999). Cultural ideologies can also work against hierarchy. For example, in both our lives, the civil rights and anticolonial movements mobilized Western arguments for equality and liberty (e.g., Klein and Licata, 2003). This aspect of SDT is compatible with much critical theory (e.g., Crenshaw et al., 1996), social representations (e.g., Moscovici, 1988), and discourse analysis (e.g., Chiapello and Fairclough, 2002) in identifying the social and political functions of ideology.

SDO and Authoritarianism

Furthermore, SDT argues that the individual's attitudes towards redistributive social ideologies, group-relevant social policies, and social groups themselves, will be strongly determined by how much one favors group-based dominance and social inequality in general. Because we have a concrete measure of SDO (see Pratto et al., 1994), one of the unique features of SDT is that it offers an empirical standard for understanding whether given cultural ideologies legitimize continued hierarchy or increased equality. [p. 427 ↓] If the desire for the establishment and maintenance of group-based social inequality and hierarchy (i.e., SDO) has a positive correlation with support for an ideology, one can regard that ideology as hierarchy-enhancing. If, on the other hand, SDO has a negative correlation with support for an ideology, one can suppose it is hierarchy-attenuating. However, a stronger test is how well an ideology *mediates* SDO and support for policies or practices that influence inequality. For example, Pratto et al. (1998) showed that *noblesse oblige*¹ was negatively correlated with SDO, and mediated between SDO and support for social welfare programs, implying that in that context, *noblesse oblige* was hierarchy-attenuating. In the same study, nationalism was

positively correlated with SDO and mediated between SDO and support for the US Gulf War against Iraq.

Although authoritarianism and SDO have both been found to be strong predictors of prejudice and hostility towards a range of groups, the two variables are also both conceptually and empirically distinct. Whereas authoritarianism was conceived from psychoanalytic theorizing as an ego defense against feelings of inadequacy and vulnerability, SDO was not conceived of as psychopathological in any sense, but merely viewed as one orientation for engaging in social life. Furthermore, except where the political system is highly unidimensional (e.g., Duriez et al., 2005), right-wing authoritarianism (RWA), and SDO are only minimally related to each other and both make strong and independent contributions to prejudice against denigrated groups such as gays, foreigners, women, Arabs, Muslims, Blacks, and Jews (e.g., Altemeyer, 1998; McFarland and Adelson, 1996).

Modern conceptualizations of right-wing authoritarianism define it as submission to ingroup authority, the social norms that these authorities endorse, and the propensity to aggress against those who are perceived as violating ingroup norms and traditions (e.g., Altemeyer, 1998). Rather than being about ingroup norms, SDO is primarily about hierarchy between groups. This conceptual distinction between RWA and SDO was recently confirmed in an experiment by Thomsen et al. (2008), who reasoned that RWA and SDO would be differentially associated with hostility against immigrants, depending upon how descriptions of immigrants were framed. High authoritarians were most hostile towards immigrants who were described as refusing to accept ingroup norms and assimilate. In contrast, high dominators were most hostile to those immigrants *who did want to* accept national norms and assimilate; thereby becoming competitors with natives (see Duckitt and Sibley, 2007; Henry et al., 2005 for empirical distinctions between RWA and SDO).

The Relation between Gender and Arbitrary Set Discrimination

Initially inspired by Jim's early discovery of consistently higher levels of racism and xenophobia among men than among women, and influenced by the biosocial analysis of Laura Betzig (1993), we began to theorize that this syndrome of greater affinity for outgroup hostility, social predation, and group-based dominance among males was most likely grounded in the notion that social dominance had slightly higher fitness-value for males than for females over the course of human evolutionary history. Thus, the "invariance hypothesis" was born, or the notion that, everything else being equal, men will tend to have higher SDO scores than women. There is now very considerable and consistent evidence in support of this hypothesis found in scores of different studies, over dozens of different cultures, and using thousands of respondents (see, for example, Pratto et al., 1997; Sidanius and Pratto, 1999; Sidanius et al., 1994b, 1995, 2000, 2006; see especially the meta-analysis of Lee et al., submitted).

Such gender differences contribute to men obtaining hierarchy-enhancing roles and to women obtaining hierarchy-attenuating roles, due not only to stereotyping, but to self-selection as well (Pratto et al., 1997; Pratto [p. 428 ↓] and Espinoza, 2001). These gender differences are not just expressed in attitudes assessed by surveys, or in hierarchy roles, but also in disproportionate acts of violence against outgroups. For example, while women sometimes participate in war, they are very rarely, if ever, the organizers and major protagonists of war (e.g., Keegan, 1993). Similarly, considering US hate crimes as an example of outgroup aggression (i.e., crimes based on arbitrary-set distinctions like race, religion, ethnicity, disability, sexual orientation), men again predominate as perpetrators, both among Whites and Blacks (Bureau of Justice Statistics, 2001).

Another important fact that we have documented concerning gender and arbitrary set discrimination is that even arbitrary set victimhood is gendered. Except for rape and child abuse, extreme violence is primarily targeted against men rather than against women. For example, White males comprised 40 percent of US hate crime victims, while White females were 25 percent of hate crime victims; Black males were 20

percent of hate crime victims, while Black females were 12 percent (Bureau of Justice Statistics, 2001). In a thorough international review of institutional discrimination, we found higher rates of victimhood among subordinate men than among subordinate women in the labor market, the retail and housing markets, the educational system, and the criminal justice system across many nations (see Sidanius and Pratto, 1999, Chapters 5–9). This gender-based asymmetry in discriminatory outcomes is called the *subordinate male target hypothesis* (SMTH; see Sidanius and Pratto, 1999; Sidanius and Veniegas, 2000).

Because this pattern of men being both the more frequent perpetrators and victims of arbitrary-set violence and discrimination is so consistent across nations, we have explored whether evolutionary theory might inform these gender differences. As mentioned above, we have argued that intergroup aggression, which is often both high risk and high gain, suits the fitness strategies of men more than those of women (e.g., Betzig, 1993).

Furthermore, we have recently begun examining this phenomenon by use of the *prepared learning paradigm*. This approach argues that conditioned fear to stimuli which have been dangerous to humans over the course of human evolutionary history (e.g., spiders and snakes) will resist extinction, while conditioned fear responses to stimuli which have not posed a threat across human evolutionary history (e.g., birds and rabbits) will be more readily extinguished (Ohman and Mineka, 2001). Applying this idea to the domain of intergroup relations, Olsson et al. (2005) used men's faces as stimuli and found that conditioned fear of facial pictures of one's racial ingroup readily extinguished, but conditioned fear of facial images of racial outgroups did not. This implies that people are "prepared" to be fearful of members of less familiar outgroups and do not easily stop fearing them.² Employing the subordinate male target hypothesis, Navarrete et al. (2009) reasoned that since outgroup males, rather than outgroup females, have posed the most lethal threats over the course of human evolutionary history, conditioned fear of outgroup faces will be most resistant to extinction when these faces are male rather than female. The experimental results were consistent with this hypothesis.

In a further extension of this reasoning, Navarrete et al., (2010) reasoned that whereas men's desire to aggress against outgroups may be motivated by dominance tendencies, women's negative reactions to outgroups may be motivated by fear of sexual coercion and rape. In fact, they found that fear extinction biases against male stimuli were predicted by aggressiveness and SDO among men, but by fear of sexual coercion among women.

These kinds of studies show what SDT has argued from its inception, namely that because some of the psychological differences between men and women are considered to be "prepared" by evolution (e.g., greater affection for the exercise of raw power, violence, and SDO among males), gender can neither be considered as just [p. 429 ↓] another form of arbitrary-set inequality, nor is gender only about sexism and irrelevant to arbitrary-set inequality. In this respect, SDT remains different from the many theories of racism that ignore some of the unique characteristics of gender relations, or consider sexism merely a parallel form of racism.³ Similarly, SDT differs from the many theories of gender that focus only on relations between men and women, and do not recognize how gender intersects with the adult-child and arbitrary set systems, nor how gender influences arbitrary set relations and social structure.⁴

Power, Not Status

Another important way that social dominance differs from most contemporary theories of intergroup relations is that SDT is centrally concerned with intergroup power, not interpersonal power, group status, minority status, intergroup contact, or other structural considerations. The heavy American focus on Black–White US relations and the minority influence school has led many theorists to focus on "minority status" and thus to ignore the examples of apartheid, Israeli treatment of Palestinians, colonization, and sexism, and other cases in which power does not merely come from numbers. Owing to the strong influence of social identity theory, with its motivational engine of positive self-regard (Tajfel and Turner, 1979), of Allport's (1954) view of prejudice as shades of disliking, and of stigma or lack of acceptance (Goffman, 1959), many other theorists have focused on the social status and social evaluation associated with groups (e.g., Fiske et al., 2002). Obviously, these intellectual cousins have made significant

contributions to intergroup relations and related processes in their own right, and our initial theorizing was strongly influenced by social identity theory (Tajfel and Turner, 1979), which we admired for linking individual and group psychological processes to social structural variables (see Sidanius et al., 2004 for an extended discussion). However, theorists before us (e.g., Ng, 1980) showed that power and status are not the same. SDT's most important epistemological assumption is that intergroup power, not which group is liked or respected more, is what matters.

Here it is important to explicate how SDT understands intergroup power. We use terms such as “dominance” and “oppression” to describe some intergroup relations, and this may lead our readers to think we are endorsing a definition of power that social psychologists rejected in the 1950s, namely that power is the ability to get another to act against his or her will, or absolute control. Because of interdependence theory (Thibaut and Kelley, 1959), and the interpersonal interaction model (e.g., Raven, 1986), many social psychologists view power as an aspect of a dyadic relation, wherein the party who can more easily exit the relationship or who exerts more influence has more power. In this view, then, power, is asymmetric interdependence (see Fiske and Berdahl, 2007 for a review). From our perspective, this influence/relational conception of power is not adequate for describing intergroup relations for three reasons.

First, many relations between groups and between group members are simply not interpersonal. There is a good deal of segregation as to where men and women work, where people of different ethnic groups and nations live, worship, and relax, and it is hard to see how not being in interpersonal intergroup contact leads to asymmetric effects for people in more and less powerful groups. What segregation in workplaces, neighborhoods, and service institutions does is to constrain which groups have access to resources, which is properly called power and not status. Another reason that intergroup relations cannot be described simply as aggregated interpersonal power relations is because discrimination is institutionalized (e.g., Feagin and Feagin, 1978). Institutional discrimination reveals that racism and sexism, for example, are not just products of asymmetric discrimination [p. 430 ↓] by individuals. There is a group-ness to intergroup relations across a society; for example, when shared categories lead to systematic differential treatment (e.g., Tilly, 1998).

Second, in addition to the kinds of interpersonal influence Raven (e.g., 1986) identified, intergroup relations may not be easily described as if there is only one kind of power. For example, Israel enjoys nationhood, a functioning society, greater military power, and greater approval by superpowers than the Palestinian people do. On the whole we would have to say that Israel is much more powerful than the Palestinians, who lack first class citizenship and statehood, have an extraordinarily high unemployment rate, receive little social recognition outside the Arab world, and are killed in high numbers by Israelis. But relatively small-scale violence by nonstate actors, like the bombing of the US Marine barracks in Lebanon in 1983, has changed policies of powerful nations (e.g., Pape, 2005). It is possible, then, that certain kinds of actions by less powerful groups, including strikes, boycotts, and nonviolent protests, and also certain kinds of violence, can effect change. To understand this we need to acknowledge that power is not untypological. In fact, having things that others desire and exerting military might may make relatively rich and powerful countries like the US and Israel vulnerable to attack by people with little money, no standing armies, and no state security. The forms of power that are relevant to intergroup relations extend beyond influence, and elaborating what these are an important agenda for intergroup relations research.

Third, a solely relational view of power does not address two important aspects of power: the extent to which people have volition or agency, and whether they can obtain basic necessities. Both the philosophy and sociology of power consider degree of freedom or *choice* to be an important aspect of power, a view also held by Lewin (1951). Having power more often enables more choice, whereas survival needs constrain some choices and necessitate others. Wellbeing and volition may not be absolute dichotomous states, such that one either has or does not have them, but they are also not relative to other people. Unlike relational views of power, SDT's assumptions about power have considered both volition and need.

SDT explicitly allows that both dominants and subordinates can have agency, but has demonstrated that groups in social hierarchies often have asymmetric outcomes. Rather than viewing power as asymmetric interdependence, though, SDT might be said to be more ontological in focusing on how the wellbeing of people in dominant and subordinate groups differs. For example, four chapters of our book (Sidanius and Pratto, 1999) review institutional discrimination to understand how basic needs like housing, income, education, and healthcare are not enjoyed as much by people in

subordinate groups as by people in dominant groups. In addition, from its beginning, SDT has shown that certain basic processes are not symmetric for people in more and less powerful groups. For example, SDO is less associated with ingroup identification for people in subordinate groups than for people in dominant groups (Sidanius et al., 1994c), and more generally, the psychological facilitators of dominance do not work as well for members of subordinate groups. Our principle of *behavioral asymmetry* argues that people in subordinate groups do not behave in ways that are as self-serving as people in dominant groups do because of their power situation. Henry (2009), in his low-status compensation theory, may be identifying part of the psychological reason this occurs.

There is another aspect of power that is implicit in SDT, but not in many other social psychological treatments of power. SDT has always acknowledged that societies often have hierarchy-attenuating individuals, cultural ideologies, and even institutions that strive against hierarchy, inequality, and exclusion. The fact that these hierarchy-attenuating forces can help the neediest have their needs met, and also effect social change, implies that SDT acknowledges the existence not just [p. 431 ↓] of *oppositional* power, but of *transformative* power. As acknowledged in social movements for empowerment (e.g., Ball, 2008) and in philosophy due to feminist theory (Wartenburg, 1990), power can be used to enable people to grow, thrive, develop, and to change relationships, not only for dominance. Indeed, dominance would be fairly easy to maintain were it not for this other kind of power. Given that we have always pointed out the importance of hierarchy-attenuating forces, it would be a mistake to assume that SDT views power only as destructive, coercive, and oppressive.

We view the recognition that social processes and the outcomes they produce are different for people in dominant versus subordinate groups to be an important and growing legacy of SDT. For example, Pratto and Espinoza (2001) tested whether job applicants of different ethnic groups and genders and who were apparently either low or high on SDO would be hired into hierarchy-enhancing or hierarchy-attenuating jobs differentially. The results showed that ethnic group moderated the previous effects found for White applicants, that men and high SDO applicants would be hired into hierarchy-enhancing jobs while women and low SDO applicants would be hired into hierarchy-attenuating jobs disproportionately (Pratto et al., 1997). Pratto and Espinoza (2001) found that Black and Hispanic applicants, regardless of their gender or SDO

levels, were placed in hierarchy-attenuating jobs over hierarchy-enhancing jobs, and that only White male applicants were sorted by their SDO level into compatible jobs. In other words, White applicants were individuated and Black and Hispanic applicants were stereotyped more in job placement. To provide a different example, Saguy et al. (2009) showed that intergroup contact is not symmetric for people in low and high power groups. High power group members prefer to talk about what they have in common with low power groups rather than the power differential, and when they do this, low power group members come to expect that power will be addressed, when in fact it won't be. One general heuristic that SDT and other group positions theories suggest is that researchers consider that group power may moderate the processes and outcomes they posit.

Social Dominance as a System

Unlike most theories in social psychology, SDT uses not just two (e.g., person–situation) but several levels of analyses. Its range of interest varies from the nature of attitudes and attitude formation at the person level (in its discussion of SDO) and the individual's construal of the social situation, to the asymmetrical behaviors of social groups, to the functions of system-wide social ideologies and the allocative decisions of social institutions (Mitchell and Sidanius, 1995). Furthermore, SDT holds that it is the interactions and intersections of these levels of analysis that account for the maintenance of social hierarchy. For example, research derived from SDT has shown that people in hierarchy-enhancing institutions tend to share the same hierarchy-enhancing legitimizing ideologies as each other (e.g., Sidanius et al., 1994a), and has also shown that performing hierarchy-enhancing roles tends to increase use of such ideologies in the discriminatory behavior of institutions (Michinov et al., 2005; Pratto et al., 1998). These kinds of intersecting processes contribute to the systematic perpetuation of hierarchy.

This example points out that SDT's assumptions that societies are social systems suggests a different kind of theorizing than is common in much of social psychology. Rather than perform critical experiments to rule out alternative explanations for large-scale outcomes like discrimination, or simply chain linear processes back in search of a root cause, SDT assumes that there is both elasticity and tenacity to interlinked social

processes. This is why we would expect the contents of legitimizing myths to change over time and to differ from culture to culture, despite the fact that the two functions they perform tend to be found [p. 432 ↓] everywhere (e.g., Pratto et al., 2000; Sidanius and Pratto, 1999).

Further, because SDT seeks to account for systematic effects, we have assumed that there are redundant processes in the system of society. For example, we assume that institutional functioning is afforded by good fit between the hierarchy-enhancing or hierarchy-attenuating character of the social institution and the attitudes and behavioral predispositions of the individuals working within these institutions. We have documented evidence for several different processes that contribute to this person–institution fit, including self-selection, hiring and attrition biases, and stereotyping (Pratto et al., 1997; Sidanius et al., 2003; van Laar et al., 1999; see review by Haley and Sidanius, 2005). This redundancy helps to make systems tenacious.

The Application of SDT to Real-World Issues

One of the strengths of SDT is its broad applicability and ability to make sense of a wide variety of intergroup phenomena and conflicts. We illustrate this wide applicability with respect to three social domains: (1) support for harsh criminal sanctions, (2) understanding the gender gap in social and political attitudes, and (3) understanding support for “terrorism” among Arab and Muslim populations.

Support for the Death Penalty

The US is among the very few countries in the world, and the only nation among the industrialized “democracies,” that still employs the death penalty.⁵ Consistent with the expectations of SDT, the evidence shows that the death penalty tends to be disproportionately used against subordinates (e.g., the poor and ethnic minorities),

especially when these subordinates have been convicted of capital crimes against dominants (see Sidanius and Pratto, 1999: 214–217).

The standard criminal justice literature suggests that Americans support the death penalty for two major reasons: (1) as a means to deter future criminality, and (2) as a means of retribution or revenge for unacceptable behavior. While we have no reason to doubt the importance of both motives as sources of death penalty support, given the fact that the death penalty is disproportionately used against subordinates rather than dominants, SDT would also expect these attitudes to serve as legitimizing ideologies in the service of continued group-based inequality. If this view is correct, we would also expect to find evidence of a substantial correlation between SDO and death penalty support, and that this relationship should be substantially mediated by both deterrence and retribution beliefs.

Evidence consistent with this view has been found using a large sample of university students and structural equation modeling. Sidanius et al. (2006) not only found that death penalty support is strongly associated with the ideologies of deterrence and retribution, but these ideologies were also found to completely mediate the positive and significant relationship between SDO and death penalty support (see [Figure 47.2](#)). Thus, alongside the other functions deterrence and retribution beliefs may serve, there is evidence that one of these functions is continued group-based inequality and dominance within American society.

Exploring the Gender Gap

Men and women have significantly different social and political attitudes and behaviors. For example, women are more likely to vote for liberal or socialist political parties, are more supportive of social welfare policies, and are less supportive of militaristic and punitive social policies than are men (see Sidanius and Pratto, 1999). The gender difference on SDO discussed earlier helps account for these differences. In extensive analyses of a wide range of social and political attitudes, Sidanius and Pratto (1999: 282–290) found [[p. 433 ↓](#)] that approximately half of the relationships between gender and these social and political attitudes could be explained in terms of the higher levels of SDO among men.

Support for Terrorism

SDT has also been applied to our understanding of support for terrorist violence against the West in general, and support of the 9/11 attack against the World Trade Center (WTC) in particular. There are at least two narratives that can be used to understand popular support for terrorist violence against the West. By far the most well-known narrative is known as the “clash of Civilizations,” thesis first proposed by Bernard Lewis (1990), and later popularized by Samuel Huntington (1993). This thesis essentially suggests that Islamic hatred of the West goes beyond mere conflicts of interest and is to be located in the wholesale rejection of Western civilization as such, not only what it does but what it is, and the principles and values that it practices and professes. These are indeed seen as innately evil, and those who promote or accept them as the “enemies of God” (Lewis, 1990).

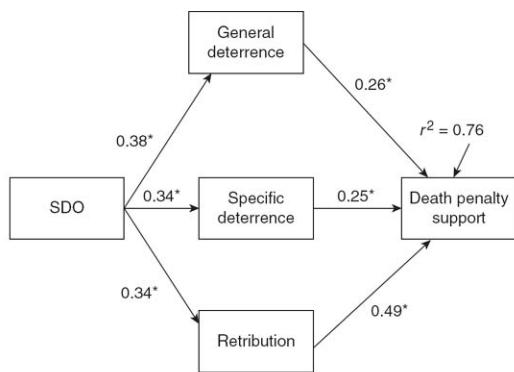
In other words, people in other cultures reject the West as being culturally degenerate and even culturally profane.

The second narrative accounts for resentment, not of Western culture, but the politics of Western dominance and hegemony. From our group dominance perspective, support for terrorism against the West can be seen as endorsement of anti- or counterdominance directed at ending the perceived oppression of the Arab and Muslims worlds by the West.

Sidanius et al. (2004) explored the relative plausibilities of these two perspectives using a sample of university students in Beirut, Lebanon. Using structural equation modeling and measures of antidominance and clash-of-civilization attributions for the attack on the WTC, Sidanius and his colleagues found that support for the 9/11 attack on the World Trade Center was strongly related to anti-dominance attributions ($r = 0.32, p < 0.05$), while being essentially unrelated to clash-of-civilization attributions ($r = -0.10, \text{n.s.}$). [p. 434 ↓] Furthermore, while a consistent body of evidence shows that support for war and anti- “terrorist” violence in the Middle-East is *positively* associated with SDO among Western populations (Crowson et al., 2006; Heaven et al., 2006; Henry et al., 2005; McFarland, 2005; Sidanius and Liu, 1992), support for “terrorist” violence against the West is *negatively* associated with SDO among Lebanese and

Middle-Eastern respondents (Henry et al., 2005). In other words, the more participants supported group-based dominance (and assumedly the present dominance of Israel and the West over Arab lands), the less one supported terrorist organizations, and the less one supported the attack on the WTC. Thus, rather than being an expression of support for group-based dominance and inequality among Lebanese students, support for terrorist violence against the West appears to be associated with counterdominance motivations. These results illustrate a central but uncommon assumption of SDT, namely that the meaning of actions and psychological states of people in dominant and subordinate groups depend on their group position.

Figure 47.2 Support for the death penalty as a function of belief in general deterrence, specific deterrence, retribution, and social dominance orientation. (From Sidanius et al. (2006))



Chi-square = 0.204, df = 1, p = 0.65, RMSEA = 0.00, p(RMSEA < 0.05) = 0.78, AGFI = 0.997

SDT's Contributions to the Intergroup Relations Literature

There are four main areas in which SDT has contributed to the social psychological literature of intergroup relations. First, in contrast to the normative emphasis on factors such as social status, self-esteem, social identity, and individual social cognition and categorization, SDT helped to reintroduce and emphasize the factor of intergroup power in both its hard forms, such as the use of institutionalized and informal physical

intimidation and violence, and soft forms such as the control of legitimizing ideologies (see Sidanius and Pratto, 1999; Mitchell and Sidanius, 1995). Consistent with realistic group conflict theory (e.g., Sherif et al., 1961), intergroup behavior is not simply driven by conflicts over social status, social regard, and symbolic rewards, but by conflict over the power to allocate social and economic resources to the benefit of one's own ingroup. We hope this has also refocused the discipline on unequal outcomes.

Second, rather than view prejudice by dominants as the only engine of inequality, SDT also emphasizes the power and agency of subordinates and their allies, for example, in hierarchy-attenuating institutions and in behavioral asymmetry. Whereas stereotyping and prejudice research often assumes only the perspective of dominant groups (e.g., Fiske et al., 2002), and stigma research focuses on the perspective of denigrated groups (e.g., Pinel, 1999), SDT not only considers both perspectives, but also how they relate to one another. This is how SDT can describe the complementarity of the behaviors of groups with different interests while showing how the actions of both sustain group dominance.

Third, although most social psychological approaches to intergroup relations limit their analyses to the intersection between the individual and the social group, SDT's analysis extends our focus from individuals to social context to institutional behavior to cultural ideologies to social structural context and reproductive patterns over historical time.

Fourth and most important, SDT puts the myriad components of intergroup beliefs, values, actions, and structure together to show how they function as a living social system. SDT has not only done this in linking processes at several different levels of social organization together, but by assuming that the stable system of hierarchy has multiple functionally redundant processes that help to stabilize it. This new way of understanding human social life may help social psychologists understand other systematic outcomes as well.

Acknowledgments

We wish to thank Arnold Ho, Nour Kteily, Jennifer Sheehy-Skeffington, and the editors [p. 435 ↓] of this volume for their valuable comments on this chapter.

Notes

1 Noblesse oblige is defined as the obligations of honorable and generous behavior of those with high rank.

2 However, for a contrary view, see Mallan et al. (2009).

3 Following Kurzban et al. (2001), we also assume that while sex may be a natural category of mind, “race” is not. Rather what we refer to as “race” may be a means of encoding coalitional alliances.

4 For a more comprehensive discussion of these issues, see Sidanius and Pratto (1999: 294–298).

5 The death penalty has been totally abolished in 46 of the 50 European nations. Abolition of the death penalty is also a condition for membership in the Council of Europe and its abolition is considered a central value to the European Union.

Jim Sidanius Felicia Pratto

References

Adorno, T.W., Frenkel Brunswik, E., Levinson, D.J. and Sanford, R.N. (1950) *The Authoritarian Personality*. New York: Norton.

Akrami, N. and Ekehammar, B. Right-wing authoritarianism and social dominance: Their roots in Big Five Personality Factors or Facets. *Journal of Individual Differences*, 27, 117–126. (2006) <http://dx.doi.org/10.1027/1614-0001.27.3.117>

Allport, G.W. (1954) *The nature of prejudice*. Menlo Park, CA: Addison-Wesley.

Altemeyer, B. (1998) The other ‘authoritarian personality’. In M.P. Zanna (ed.), *Advances in Experimental Social Psychology*, 30, 47–92. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960382-2>

Ball, J.B. A second look at the industrial areas foundation: Lessons for Catholic social thought and ministry . *Horizons* , 35, 271–297. (2008)

Beshoar, B.B. (1957) Out of the Depths: The Story of John R. Lawson, a Labor Leader . Denver: Colorado Historical Commission and Denver Trades and Labor Assembly.

Betzig, L. (1993) Sex, succession, and stratification in the first six civilizations: How powerful men reproduced, passed power on to their sons, and used power to defend their wealth, women and children . In L. Ellis (ed.), *Social Stratification and Socio economic Inequality: A Comparative Biosocial Analysis* , pp. 37–74. New York: Praeger.

Blair, I.V. and Banaji, M.R. Automatic and controlled processing in stereotype priming . *Journal of Personality and Social Psychology* , 70, 1142–1163. (1996) <http://dx.doi.org/10.1037/0022-3514.70.6.1142>

Blumer, H. Race prejudice as a sense of group position . *Pacific Sociological Review* , 1, 3–5. (1960)

Bretz, R.D. and Judge, T.A. Person-organization fit and the theory of work adjustment: Implications for satisfaction, tenure, and career success . *Journal of Vocational Behavior* , 44, 32–54. (1994) <http://dx.doi.org/10.1006/jvbe.1994.1003>

Bureau of Justice Statistics (2001) Hate crimes reported in NIBRS, 1997–1999 . Available at: <http://www.ojp.usdoj.gov/bjs/pub/pdf/hcrn99.pdf>.

Chiapello, E. and Fairclough, N. Understanding the new management ideology: A transdisciplinary contribution from critical discourse analysis and new sociology of capitalism . *Discourse and Society* , 13, 185–208. (2002) <http://dx.doi.org/10.1177/0957926502013002406>

Clark, K.B. and Clark, M.P. (1947) Racial identification and preference in Negro children . In T.M. Newcomb, ed. and E.L. Hartley (eds), *Readings in Social Psychology* , pp. 169–178. New York: Holt, Rinehart, and Winston.

Crenshaw, K., Gotanda, N., Peller, G. and Thomas, K. (1996) *Critical Race Theory: The Key Writings That Formed the Movement* . New York: The New Press.

Crowson, H.M., DeBacker, K. and Thomas, S.J. The role of authoritarianism, perceived threat, and need for closure or structure in predicting post 9/11 attitudes and beliefs . The Journal of Social Psychology , 146, 733–750. (2006) <http://dx.doi.org/10.3200/SOCP.146.6.733-750>

Dahl, R.A. (1989) Democracy and its Critics . New Haven: Yale University Press.

Devine, P.G. Stereotypes and prejudice: Their automatic and controlled components . Journal of Personality and Social Psychology , 56, 5–18. (1989) <http://dx.doi.org/10.1037/0022-3514.56.1.5>

Dickemann, M. (1979) Female infanticide, reproductive strategies, and social stratification: A preliminary model . In N.A. Chagnon, ed. and W. Irons (eds), Evolutionary Biology and Human Social Behavior: An Anthropological Perspective , pp. 321–367. North Scituate, MA: Duxbury Press.

Duckitt, J. and Sibley, C.G. Right-wing authoritarianism, social dominance orientation and the dimensions of generalized prejudice . European Journal of Personality , 21, 113–130. (2007) <http://dx.doi.org/10.1002/per.614>

Duriez, B., Van Hiel, A. and Kossowska, M. Authoritarianism and social dominance in Western and Eastern Europe: The importance of the sociopolitical context and political interest and involvement . Political Psychology , 26, 299–320. (2005) <http://dx.doi.org/10.1111/j.1467-9221.2005.00419.x>

Ekehammar, B. Sex differences in socio political attitudes revisited . Educational Studies , 11, 3–9. (1985) <http://dx.doi.org/10.1080/0305569850110101>

Ekehammar, B. and Sidanius, J. Sex differences in socio-political ideology: A replication and extension . British Journal of Social Psychology , 21, 249–257. (1982) <http://dx.doi.org/10.1111/j.2044-8309.1982.tb00546.x>

Feagin, J.R. and Feagin, C.B. (1978) Discrimination American Style: Institutional Racism and Sexism . Englewood Cliffs, Prentice Hall.

Fiske, S.T. and Berdahl, J. (2007) Social power . In A.W. Kruglanski, ed. and E.T. Higgins (eds), *Social Psychology: Handbook of Basic Principles* , 2nd Edition, pp. 678–692. New York: Guilford Press.

Fiske, S.T., Cuddy, A.J.C., Glick, P. and Xu, J. A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition . *Journal of Personality and Social Psychology* , 82, 878–902. (2002) <http://dx.doi.org/10.1037/0022-3514.82.6.878>

Furnham, A. Adolescents' sociopolitical attitudes: A study of sex and national differences . *Political Psychology* , 6, 621–636. (1985) <http://dx.doi.org/10.2307/3791020>

Gaertner, S.L. and Dovidio, J.F. (1986) The aversive form of racism . In S.L. Gaertner, ed. and J.F. Dovidio (eds), *Prejudice, Discrimination, and Racism: Theory and Research* , pp. 61–89, NY: Academic Press.

Goffman, E. (1959) *The presentation of self in everyday life* . Garden City, NY: Doubleday.

Gramsci, A. (1971). Selections from the Prison Notebooks . London: Wishart.

Haley, H. and Sidanius, J. Person-organization congruence and the maintenance of group-based social hierarchy: A social dominance perspective . *Group Processes & Intergroup Relations* , 8, 187–203. (2005) <http://dx.doi.org/10.1177/1368430205051067>

Hamilton, D.L. and Gifford, G.K. Illusory correlation in interpersonal perception: A cognitive basis of stereotype judgments . *Journal of Experiment Social Psychology* , 12, 392–407. (1976) <http://dx.doi.org/10.1016/S0022-1031%2876%2980006-6>

Heaven, P.C.L., Organ, L., Supavadeeprasit, S. and Leeson, P. War and prejudice: A study of social values, right-wing authoritarianism, and social dominance orientation . *Personality and Individual Differences* , 40, 599–608. (2006) <http://dx.doi.org/10.1016/j.paid.2005.08.005>

Henry, P.J. Low-status compensation: A theory for understanding the role of status in cultures of honor . *Journal of Personality and Social Psychology* , 97, 451–466. (2009) <http://dx.doi.org/10.1037/a0015476>

Henry, P.J., Sidanius, J., Levin, S. and Pratto, F. Social dominance orientation, authoritarianism, and support for intergroup violence between the Middle East and America . *Political Psychology* , 26, 569–583. (2005) <http://dx.doi.org/10.1111/j.1467-9221.2005.00432.x>

Hood, R. and Cordovil, G. (1992) Race and Sentencing: A Study in the Crown Court. A Report for the Commission for Racial Equality . Oxford: Clarendon Press.

Huntington, S. The clash of civilizations . *Foreign Affairs* , 72, 22–49. (1993) <http://dx.doi.org/10.2307/20045621>

Jost, J.T. and Banaji, M.R. The role of stereotyping in system-justification and the production of false consciousness . *British Journal of Social Psychology* , 33, 1–27. (1994) <http://dx.doi.org/10.1111/j.2044-8309.1994.tb01008.x>

Katz, I. and Hass, R.G. Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures . *Journal of Personality and Social Psychology* , 55, 893–905. (1988) <http://dx.doi.org/10.1037/0022-3514.55.6.893>

Keegan, J. (1993) *The History of Warfare* . New York: Alfred A. Knopf.

Klein, O. and Licata, L. When group representations serve social change: The speeches of Patrice Lumumba during the Congolese decolonization . *British Journal of Social Psychology* , 42, 571–593. (2003) <http://dx.doi.org/10.1348/014466603322595284>

Kurzban, R., Tooby, J. and Cosmides, L. Can race be erased? Coalitional computation and social categorization . *Proceedings of the National Academy of Science* , 98, 15387–15392. (2001) <http://dx.doi.org/10.1073/pnas.251541498>

Lee, I., Pratto, F. and Johnson, B.T. (submitted) Intergroup Consensus/Disagreement in Support of Group-Based Hierarchy: An Examination of Socio-Structural and Psycho-Cultural Factors .

Lenski, G.E. (1984) Power and Privilege: A Theory of Social Stratification . Chapel Hill: University of North Carolina Press.

Lewin, K. (1951) Field theory in social science: Selected theoretical papers . (D. Cartwright, ed.). New York: Harpers & Brothers.

Lewis, B. The roots of Muslim rage: Why so many Muslims deeply resent the West, and why their bitterness will not easily be mollified . The Atlantic Monthly , 266, 47–60. (1990)

Mallan, K.M., Sax, J. and Lipp, O.V. Verbal instructions abolishes fear conditioned to racial out-group faces . Journal of Experimental Social Psychology , 45, 1303–1307. (2009) <http://dx.doi.org/10.1016/j.jesp.2009.08.001>

Marjoribanks, K. Sex related differences in socio political attitudes: A replication . Educational Studies , 7, 1–6. (1981) <http://dx.doi.org/10.1080/0305569810070101>

Marx, K. and Engels, F. (1846) The German Ideology, (1970) . New York: International Publishers.

McFarland, S.G. On the eve of war: Authoritarianism, social dominance, and American students' attitudes toward attacking Iraq . Personality and Social Psychology Bulletin , 31, 360–367. (2005) <http://dx.doi.org/10.1177/0146167204271596>

McFarland, S.G. and Adelson, S. (1996) An omnibus study of personality, values and prejudices . Paper presented at the Annual Convention of the International Society for Political Psychology, Vancouver, British Columbia, July 1996.

Michels, R. (1911) Political Parties: A Sociological Study of the Oligarchical Tendencies of Modern Democracy, (1962) . New York: Free Press.

Michinov, N., Dambrun, M., Guimond, S. and Meot, A. Social dominance orientation, prejudice, and discrimination: A new computer method for studying discrimination behavior . Behavioral Research Methods , 37, 91–98. (2005) <http://dx.doi.org/10.3758/BF03206402>

Mitchell, M. and Sidanius, J. Social hierarchy and executions in the United States: A social dominance perspective. *The death penalty: A social dominance perspective*. *Political Psychology*, 16, 591–619. (1995) <http://dx.doi.org/10.2307/3792228>

Mosca, G. (1896) *The Ruling Class: Elements of Political Science*, Edition (1939). New York: McGraw-Hill.

Moscovici, S. Notes toward descriptions of social representations. *European Journal of Social Psychology*, 18, 211–250. (1988) <http://dx.doi.org/10.1002/ejsp.2420180303>

Navarrete, C.D., McDonald, M.M., Molina, L.M. and Sidanius, J. The psychological architecture of race and gender bias: An outgroup male target hypothesis. *Journal of Personality and Social Psychology*, 98, 933–945. (2010) <http://dx.doi.org/10.1037/a0017931>

Navarrete, C.D., Olsson, A., Ho, A.K., Mendes, W.B., Thomsen, L. and Sidanius, J. Fear extinction to an outgroup face: The role of target gender. *Psychological Science*, 20, 155–158. (2009) <http://dx.doi.org/10.1111/j.1467-9280.2009.02273.x>

Ng, S.H. (1980) *The Social Psychology of Power*. NY: Academic Press.

Ohman, A. and Mineka, S. Fears, phobias and preparedness: Towards an evolved model of fear and learning. *Psychological Review*, 108, 483–522. (2001). <http://dx.doi.org/10.1037/0033-295X.108.3.483>

Olsson, A., Ebert, J.P., Banaji, M.R. and Phelps, E.A. The role of social groups in the persistence of learned fear. *Science*, 309, 785–787. (2005) <http://dx.doi.org/10.1126/science.1113551>

Pape, R. (2005) *Dying to Win: The Strategic Logic of Suicide Terrorism*. New York: Random House.

Pareto, V. (1901) *The Rise and Fall of the Elites*, (1979). New York: Arno Press.

Pareto, V. (1935) *The Mind and Society: A Treatise on General Sociology*, (1963). New York: Dover.

Pinel, E.C. Stigma consciousness: The psychological legacy of social stereotypes . Journal of Personality and Social Psychology , 76, 114–128. (1999) <http://dx.doi.org/10.1037/0022-3514.76.1.114>

Pratto, F. (1999) Social dominance theory . In M.P. Zanna (ed.), Advances in Experimental Social Psychology , 31, 192–263. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960274-9>

Pratto, F. and Espinoza, P. Gender, ethnicity and power (2001) . Journal of Social Issues , 57, 763–780. (2001) <http://dx.doi.org/10.1111/0022-4537.00240>

Pratto, F., Liu, J.H., Levin, S., Sidanius, J., Shih, M., Bachrach, H. and Hegarty, P. Social dominance orientation and the legitimization of inequality across cultures . Journal of Cross-Cultural Psychology , 31, 369–409. (2000) <http://dx.doi.org/10.1177/0022022100031003005>

Pratto, F., Sidanius, J. and Levin, S. Social dominance theory and the dynamics of intergroup relations: Taking stock and looking forward . European Review of Social Psychology , 17, 271–320. (2006) <http://dx.doi.org/10.1080/10463280601055772>

Pratto, F., Sidanius, J., Stallworth, L.M. and Malle, B.F. Social dominance orientation: A personality variable predicting social and political attitudes . Journal of Personality and Social Psychology , 67, 741–763. (1994) <http://dx.doi.org/10.1037/0022-3514.67.4.741>

Pratto, F., Stallworth, L.M. and Conway-Lanz, S. Social dominance orientation and the legitimization of policy . Journal of Applied Social Psychology , 28, 1853–1875. (1998) <http://dx.doi.org/10.1111/j.1559-1816.1998.tb01349.x>

Pratto, F., Stallworth, L., Sidanius, J. and Siers, B. The gender gap in occupational role attainment: A social dominance approach . Journal of Personality and Social Psychology , 72, 37–53. (1997) <http://dx.doi.org/10.1037/0022-3514.72.1.37>

Raven, B.H. A taxonomy of power in human relations . Annals of Psychiatry , 16, 633–636. (1986)

Rokeach, M. (1973) The Nature of Human Values . New York: Free Press.

Saguy, T., Tasuch, N., Dovidio, J.F. and Pratto, F. The irony of harmony: intergroup contact can produce false expectations for group equality . Psychological Science , 20, 114–121. (2009) <http://dx.doi.org/10.1111/j.1467-9280.2008.02261.x>

Sanday, P. (1981) Female Power and Male Dominance . Cambridge: Cambridge University Press.

Sherif, M., Harvey, O.J., White, B.J., Hood, W.R. and Sherif, C. (1961) Intergroup Conflict and Cooperation: The Robbers' Cave Experiment . Norman, OK: Institute of Group Relations, University of Oklahoma.

Sidanius, J., Devereux, E. and Pratto, F. A comparison of symbolic racism theory and social dominance theory: Explanations for Racial Policy Attitudes . Journal of Social Psychology , 132, 377–395. (1992) <http://dx.doi.org/10.1080/00224545.1992.9924713>

Sidanius, J. and Ekehammar, B. Sex-related differences in socio-political ideology . Scandinavian Journal of Psychology , 21, 17–26. (1980). <http://dx.doi.org/10.1111/j.1467-9450.1980.tb00336.x>

Sidanius, J. and Ekehammar, B. Sex, political party preference and higher-order dimensions of socio-political ideology . Journal of Psychology , 115, 233–239. (1983). <http://dx.doi.org/10.1080/00223980.1983.9915440>

Sidanius, J., Henry, P.J., Pratto, F. and Levin, S. Arab attributions for the attack on America: The case of Lebanese sub-elites . Journal of Cross-Cultural Psychology , 35, 403–416. (2004) <http://dx.doi.org/10.1177/0022022104266106>

Sidanius, J., Levin, S., Liu, J.H. and Pratto, F. Social dominance orientation and the political psychology of gender: an extension and cross-cultural replication . European Journal of Social Psychology , 30, 41–67. (2000) <http://dx.doi.org/10.1002/%28SICI%291099-0992%28200001/02%2930:1%3C41::AID-EJSP976%3E3.0.CO;2-O>

Sidanius, J. and Liu, J. Racism, support for the Persian Gulf War, and the police beating of Rodney King: A social dominance perspective . Journal of Social Psychology , 132, 685–700. (1992) <http://dx.doi.org/10.1080/00224545.1992.9712099>

Sidanius, J., Liu, J., Shaw, J. and Pratto, F. Social dominance orientation, hierarchy-attenuators and hierarchy-enhancers: Social dominance theory and the criminal justice system . *Journal of Applied Social Psychology* , 24, 338–366. (1994a) <http://dx.doi.org/10.1111/j.1559-1816.1994.tb00586.x>

Sidanius, J., Mitchell, M., Haley, H. and Navarrete, C.D. Support for harsh criminal sanctions and criminal justice beliefs: A social dominance perspective . *Social Justice Research* , 19, 433–449. (2006) <http://dx.doi.org/10.1007/s11211-006-0026-4>

Sidanius, J. and Pratto, F. (1993) The inevitability of oppression and the dynamics of social dominance . In P. Sniderman, ed. and P. Tetlock (eds), *Prejudice, Politics, and the American Dilemma* , pp. 173–211. Stanford: Stanford University Press.

Sidanius, J. and Pratto, F. (1999) *Social Dominance: An Intergroup Theory of Social Hierarchy and Oppression* . Cambridge, MA: Cambridge University Press.

Sidanius, J., Pratto, F. and Bobo, L. Social dominance orientation and the political psychology of gender: A case of invariance? *Journal of Personality and Social Psychology* , 67, 998–1011. (1994b) <http://dx.doi.org/10.1037/0022-3514.67.6.998>

Sidanius, J., Pratto, F. and Brief, D. Group dominance and the political psychology of gender: A cross-cultural comparison . *Political Psychology* , 16, 381–396. (1995) <http://dx.doi.org/10.2307/3791836>

Sidanius, J., Pratto, F. and Rabinowitz, J. Gender, ethnic status, ingroup attachment and social dominance orientation . *Journal of Cross-Cultural Psychology* , 25, 194–216. (1994c) <http://dx.doi.org/10.1177/0022022194252003>

Sidanius, J., Pratto, F., Sinclair, S. and van Laar, C. Mother Teresa Meets Genghis Khan: The dialectics of hierarchy-enhancing and hierarchy-attenuating career choices . *Social Justice Research* , 9, 145–170. (1996) <http://dx.doi.org/10.1007/BF02198077>

Sidanius, J., Sinclair, S. and Pratto, F. Social dominance orientation, gender and increasing college exposure . *Journal of Applied Social Psychology* , 36, 1640–1653. (2006) <http://dx.doi.org/10.1111/j.0021-9029.2006.00074.x>

Sidanius, J., Van Laar, C., Levin, S. and Sinclair, S. Social hierarchy maintenance and assortment into social roles: a social dominance perspective . Group Processes and Intergroup Relations , 6, 333–352. (2003) <http://dx.doi.org/10.1177/13684302030064002>

Sidanius, J. and Veniegas, R.C. (2000) Gender and race discrimination: The interactive nature of disadvantage . In S. Oskamp (ed.), Reducing Prejudice and Discrimination the Claremont Symposium on Applied Social Psychology pp. 47–69. Mahwah, NJ: Lawrence Erlbaum Associates.

Stephan, W.G. and Finlay, K. The role of empathy in improving intergroup relations . Journal of Social Issues , 55, 729–743. (1999) <http://dx.doi.org/10.1111/0022-4537.00144>

Tajfel, H. and Turner, J.C. (1979). An integrative theory of intergroup conflict . In W.G. Austin, ed. and S. Worchel (eds), The Social Psychology of Intergroup Relations , pp. 33–47. Monterey, CA: Brooks/Cole.

Tajfel, H. and Turner, J.C. (1986) The social identity theory of intergroup behavior . In S. Worchel, ed. and W.G. Austin (eds), Psychology of Intergroup Relations , pp. 7–24. Chicago: Nelson-Hall.

Thibaut, J.W. and Kelley, H.H. (1959) The Social Psychology of Groups . New York: Wiley.

Thomsen, L., Green, E.G.T. and Sidanius, J. We will hunt them down: How social dominance orientation and right-wing authoritarianism fuel ethnic persecution of immigrants in fundamentally different ways . Journal of Experimental Social Psychology , 44, 1455–1464. (2008) <http://dx.doi.org/10.1016/j.jesp.2008.06.011>

Tilly, C. (1998) Durable Inequality . Berkeley, CA: University of California Press.

Trivers, R. (1972) Parental investment and sexual selection . In B. Campbell (ed.), Sexual Selection and the Descent of Man , pp. 136–179. New York: Aldine de Gruyter.

van den Berghe, P.L. (1978) *Man in Society: A Biosocial View* . New York: Elsevier North Holland.

van Laar, C., Sidanius, J., Rabinowitz, J. and Sinclair, S. The three R's of academic achievement: Reading, 'Riting, and Racism . *Personality and Social Psychology Bulletin* . 25, 139–151. (1999) <http://dx.doi.org/10.1177/0146167299025002001>

Wartenberg, T.E. (1990). *The Forms of Power: From Domination to Transformation* . Philadelphia: Temple University Press.

<http://dx.doi.org/10.4135/9781446249222.n47>

Handbook of Theories of Social Psychology

The Common Ingroup Identity Model

Contributors: Samuel L. Gaettner & John F. Dovidio

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "The Common Ingroup Identity Model"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n48>

Print pages: 439-458

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n48>

[p. 439 ↓]

Chapter 48: The Common Ingroup Identity Model

Abstract

This chapter reviews, from both a theoretical and personal perspective, the development of the common ingroup identity model, a social categorization-based perspective for reducing intergroup bias and improving intergroup relations. The model proposes that inducing members of different groups conceive of themselves as belonging to the same, more inclusive entity produces more positive beliefs, feelings, and behaviors toward one another. The chapter demonstrates the metamorphosis of our research interests from identifying a problem, aversive racism, to addressing the issue, the common ingroup identity model. First, we describe our personal journey leading to the development of the common ingroup identity model. Second, we discuss the model's historical intellectual roots and the empirical support for this point of view. Third, we consider how the model can be applied to address actual social problems, including prejudice, discrimination, and racism. The general message that this chapter conveys is that it is not only empirical support that advances theory but also the practical and conceptual challenges to the theory that stimulate conceptual development and new perspectives on enduring social problems.

Introduction

We have never been just a collection of individuals or a collection of red states and blue states; we are and always will be the United States of America.

Barack Obama (Election Eve, November 4, 2008).

This chapter is about the development of the common ingroup identity model (Gaertner and Dovidio, 2000), a social categorization-based perspective for reducing intergroup bias and improving intergroup relations. Fundamentally, this model proposes that if members of different groups (e.g., red states and blue states) would conceive of themselves as belonging to the same more inclusive entity (i.e., the United States, as Barack Obama's quote declares), then they would have more positive beliefs, feelings, and behaviors toward one another. Thus, factors that encourage members of different groups to think in more inclusive ways (e.g., in terms of "we," "us," or "our") can promote more harmonious intergroup relations.

[p. 440 ↓]

Herein, we describe our personal journey leading to the development of the common ingroup identity model. We identify critical junctures, which are much clearer in hindsight than as we charted our paths, along the route leading to our attraction to this perspective. Then we discuss the model's historical intellectual roots and the empirical support for this point of view. After that, we consider how the model can be applied to address actual social problems, including prejudice, discrimination, and racism, which is where our journey began. We conclude with a discussion of directions that we and others have pursued as the evidence revealed that the theory's initial assumptions, although useful, were incomplete or more complex than we initially presumed and therefore present several problems and challenges for future research.

Our Personal Journey

We trace the development of the common ingroup identity model back over 30 years. Various seeds of the ideas that formed the basis of the model appeared in our studies of race and helping (e.g., Gaertner and Dovidio, 1977) and prosocial behavior more generally (Piliavin et al., 1981), as well as in our research about identifying and combating subtle racism (Gaertner and Dovidio, 1986). In the concluding chapter of our edited volume, *Prejudice, Discrimination, and Racism*, we wrote, "[T]he research challenge is to discover techniques and strategies that induce members of separate groups to conceive of the aggregate as one entity, and then to examine

whether this perception facilitates cooperativeness, acceptance, and personalized interactions" (Gaertner and Dovidio, 1986: 326).

The ideas coalesced into the initial formal presentation of the theoretical framework in 1993 (Gaertner et al., 1993), which was published in the *European Review of Social Psychology*. The most elaborate presentation of the model appeared in the monograph, *Reducing Intergroup Bias: The Common Ingroup Identity Model* (Gaertner and Dovidio, 2000). In addition, we continue to refine, revise, and update the model (e.g., Dovidio et al., 2008). Consistent with the goals articulated by the editors of this volume, the present account presents our personal story relating to the development of the model and sets its key assumptions into their proper intellectual context.

Our journey is a collaborative one. The common ingroup identity model developed out of a collaborative relationship between the authors that began almost 36 years ago and our early research, together and apart, sowed the seeds for the major theme of the model; namely, the cognitive and motivational processes initiated by the recognition of ingroup membership – that is, "we-ness."

Sam Gaertner, an assistant professor at the University of Delaware and a new PhD from the City University of New York Graduate Center, had been working on the topic of aversive racism, a contemporary and subtle, but insidious, manifestation of racial bias. His work on this topic began with his PhD dissertation that obtained a serendipitous and provocative finding. A study in the dissertation (Gaertner, 1973) involved the willingness of registered Liberal and Conservative Party members in New York City to help a Black or White motorist whose car had broken down on a local highway. Confederates, who were identifiable as Black or White on the basis of their dialects, made telephone calls, claiming to have been dialing their mechanic's number from a public telephone along the highway. The callers explained that they now needed the respondent's help to call a mechanic because they used their last coin for this wrong-number call.

Consistent with previous research using paper and pencil measures of political ideology and racial attitudes (Adorno et al., 1950), Conservative Party members discriminated by helping Black callers less frequently than White callers, whereas Liberal Party members did not discriminate in terms of helping. Surprisingly, however, Liberal Party members discriminated in a different way. [p. 441 ↓] Although Liberals helped Black and White

callers equivalently when they knew their assistance was needed, they terminated this encounter more readily for Black than for White callers *prior to* learning fully of the caller's need for *their* help. These latter results were initially puzzling but became more understandable in the context of the notion of aversive racism (Kovel, 1970). This perspective suggested that liberals may be unconsciously biased and engage in subtle rather than blatant discrimination.

In 1973, Jack Dovidio, following the advice of his undergraduate advisors, came to Delaware as a graduate student after earning his BA at Dartmouth. He came with a range of interests, including altruism and nonverbal behavior, but with a primary personal interest in intergroup relations and prejudice. After about two years, our own group-based boundaries as faculty advisor and graduate student disintegrated and we became a research team and close personal friends ever since.

As an undergraduate, together with his advisor Bill Morris, Jack investigated the implications of an idea proposed by LeVine and Campbell (1972) that similarity of fate (see also Campbell, 1958) regarding a highly threatening event increases awareness of ingroup identity and the magnification of positive behaviors (e.g., prosocial action) that accompany ingroup membership. In a study by Dovidio and Morris (1975), while the participant and a confederate partner were waiting to participate in either the same or different experiments involving stressful electric shocks or nonstressful word association tasks, the confederate "accidentally" knocked a container of 100 pencils to the floor. The results revealed that prosocial behavior (involving the percentage of participants who helped or the number pencils participants picked up) was highest when participants expected to participate in the same highly stressful experiment than when they were to participate in the same low stress experiment or when they were to participate in different experiments. Dovidio and Morris concluded that "facing stress together increases the likelihood of positive behaviors such as helping, presumably, by making the ingroup–outgroup distinction more salient and increasing 'we feelings'" (1975: 148).

In 1974, with support from the Office of Naval Research, we began a series of studies to explore predictions derived from the aversive racism framework. The fundamental premise of this research on aversive racism was that many Whites who consciously support egalitarian principles, endorse a liberal political ideology, and believe themselves to be non-prejudiced, also harbor negative attitudes about Blacks and other

historically disadvantaged groups. These unconscious negative feelings and beliefs develop as a consequence of normal, almost unavoidable, and frequently functional cognitive, motivational, and social-cultural processes. As a consequence, whereas old-fashioned racists exhibit a direct and overt pattern of discrimination, aversive racists' actions may appear more variable and inconsistent. At times they discriminate (manifesting their negative feelings), and at other times they do not (reflecting their egalitarian beliefs). Specifically, we hypothesized that aversive racists discriminate against Blacks mainly (a) when norms in a situation are weak or ambiguous, so that it is difficult for aversive racists to judge their behavior as inappropriate; or (b) when they can justify or rationalize their negative behavior on the basis of some factor other than race. Thus discrimination occurs without challenging their nonprejudiced self-image. For example, the appropriateness of Liberals terminating their encounter with the Black motorist before learning that *their* help was necessary has no socially prescribed answer. Thus, Liberals discriminated only when their behavior could not be condemned, by others or themselves, as inappropriate.

Over the next decade, we conducted a number of collaborative projects studying race and helping, which not only provided support for the aversive racism framework (Dovidio and Gaertner, 2004; Gaertner and Dovidio, 1986) but also contributed to an emerging literature on bystander intervention. Much of our data were consistent with a [p. 442 ↓] general model of emergency intervention introduced by Jane Piliavin and Irving Piliavin (1973). Our convergence of findings and interests led to a collaboration with Jane Piliavin and Russ Clark, resulting in the book, *Emergency Intervention* (Piliavin et al., 1981), which attempted to explain *why* bystanders might intervene into the problems of other people. In this volume, we collectively elaborated and revised the Piliavin and Piliavin model, which proposed that bystander intervention was motivated by a desire to reduce unpleasant arousal elicited by witnessing an emergency while weighing the rewards and costs associated with various alternative actions for reducing this arousal. In the revised *arousal: cost-reward model*, "we-ness" between the bystander and the victim was identified as playing a central role in influencing *both* arousal and the perceptions of costs and rewards of the alternative actions, and thus ultimately was a critical factor in a bystander's responsiveness to an emergency.

In hindsight, we regard the *emergency intervention* project as especially critical in shaping our theoretical perspective and creating a conceptual foundation for the

common ingroup identity model. Over the three-year period that we worked on revising the arousal: cost–reward model, our focus became riveted on the value of “we-ness,” particularly among virtual strangers, in a context in which one of them needed the assistance of the other. Our interest in the importance of the social connection between individuals expanded our perspective to consider more fully the nature of *intergroup* processes, beyond the intrapsychic processes, such as the ambivalence between the conscious egalitarian values and unconscious bias, we had been focusing on in the study of aversive racism.

Intellectual Context of the Theory

Our consideration of the value of “we-ness” in the study of prosocial behavior significantly altered and widened our perspective on processes underlying racism and on potential ways to combat contemporary forms of racism such as aversive racism. In this section, we describe how changes in the theoretical landscape, influenced by European social psychologists, shaped our perspective, discuss the profound effects of social categorization, and describe the model and present the basic evidence supporting it.

The Changing Theoretical Landscape

As a direct consequence of focusing on “we-ness,” we became very interested in the European perspective on the importance of group membership and social identity for influencing intergroup attitudes and the behavior of individuals more generally. Our research on aversive racism and our general perspective on social psychology up until this point had been shaped primarily by North American social psychology, which had a strong emphasis in the individual (Steiner, 1974). The dominant theories of prejudice at that time conceived of bias as an attitude that was shaped by a number of individual-level processes. It was hypothesized to originate from socialization experiences with punitive parents who supported hierarchical relations (Adorno et al., 1950) and personal frustration (Dollard et al., 1939). These theories focused on individuals motivated to satisfy their own individual needs. The social–cognitive movement of the 1970s

further implicated intrapersonal processes related to the types of heuristics that people use when thinking about a complex social world, and how these cognitive short-cuts contribute to stereotyping and prejudice (see Fiske and Taylor, 1991).

In contrast to the North American movement in social psychology toward more microlevel mechanisms in prejudice, European social psychology placed more emphasis on prejudice as an intergroup process, involving mechanisms such as collective identity that emphasized the distinct relationships that [p. 443 ↓] people have with members of their own group and other groups. European social psychology reflected the influence of Gestalt psychology, in which the focus included the individual but within a dynamic social field in which the individual is an integral part. Similar to the European approach, Sherif et al. (1961), in their classic Robbers Cave studies, emphasized the foundational role of social categorization in demarcating ingroup and outgroup membership and proposed that the functional relations between ingroups and outgroups are critical in determining intergroup attitudes. According to this position, competition between groups produces prejudice and discrimination, whereas intergroup interdependence and cooperative interaction that result in successful outcomes reduce intergroup bias.

In the tradition of European social psychology, social identity theory (Tajfel and Turner, 1979) and self-categorization theory (Turner, 1985) view the distinction between personal identity and social identity as a critical one because of its potential transformational consequences. What struck us as especially intriguing was the notion emerging from this literature that “the attractiveness of an individual is not constant, but varies with ingroup membership” (Turner, 1985: 60). Work in this vein that was particularly influential to us was Brewer's (1979) “cognitive–motivational analysis” of ingroup bias in the minimal intergroup situation. Her qualitative review of research revealed that features that increased the salience of ingroup–outgroup categorization (e.g., competition, similarity) were associated with increased intergroup bias. In particular, Brewer's analysis suggested that increases in intergroup bias were more often due to increased positive regard for ingroup members rather than to the devaluation of outgroup members. Brewer proposed, “Reconceptualizing the process of intergroup differentiation tends to shift the focus of attention from the negative implications of out-group perceptions to the positive consequences of in-group formation” (1979: 322). According to Brewer, upon group formation, ingroup members

are moved closer to and become less differentiated from the self and consequently are accorded more positive beliefs, feelings, and behaviors.

The possibility that intergroup bias could reflect positive orientations toward ingroup members rather than negative orientations toward outgroup members led us to consider the appropriateness of this perspective for understanding aversive racism. Whereas our focus had been primarily on how unconscious negative racial attitudes drive the subtle discrimination we observed, we recognized that discrimination among aversive racists may also be motivated by a tendency to behave in especially positive ways toward Whites. We thought that ingroup favoritism effects might help explain why aversive racists do not report possessing negative racial affect and genuinely regard themselves as nonprejudiced: Their actions may not be based primarily on negative orientations toward Blacks but rather on positive feelings about other ingroup members, Whites (Gaertner et al., 1997). Nevertheless, although positive ingroup regard may not be regarded as prejudice, it is important to recognize that its consequences may often be just as pernicious as anti-outgroup sentiment.

This insight was illuminating when considering ways to combat aversive racism. Over the years, we frequently presented to general as well as to professional audiences alerting them to the existence and dangers of aversive racism. After these presentations, people often asked, if people are unaware of their prejudice, how can we change their attitudes and, more importantly, their behavior toward Blacks? The research on ingroup–outgroup categorization and ingroup favoritism provided an essential clue. If aversive racists could conceive of Blacks and Whites primarily within a common, shared identity (e.g., employees of the same organization, citizens of the same nation) instead of two different racial groups, then the forces of ingroup favoritism could be extended to create more positive orientations toward Blacks.

[p. 444 ↓]

We were far from being alone in recognizing the potential central role of social categorization in intergroup relations. Allport (1954) in his seminal work, *The Nature of Prejudice*, wrote about the normality and inevitability of social categorization: “The human mind must think with the aid of categories ... Categories are the basis for normal prejudgment. We cannot possibly avoid this process” (1954: 20). In his highly influential

article, "Cognitive aspects of prejudice," Tajfel (1969) further emphasized that the cognitive bases of prejudice were not primarily irrational or psychopathological but rather directly related to social categorization and the search for social meaning. With this in mind, it might be easier to understand our attraction to the common ingroup identity model and the potential value of ingroup membership and the recognition of "we-ness" across group lines. Therefore, we began to seek a remedy for aversive racism that reduces indifference and increases the perceptions of connectedness between people across group lines: inducing people to recategorize others as members of a common ingroup.

Social Categorization and the Benefits to Ingroup Members

When people or objects are categorized into groups, actual differences between members of the same category tend to be perceptually minimized (Tajfel, 1969) and often ignored in making decisions or forming impressions. Members of the same category seem to be more similar than they actually are, and more similar than they were before they were categorized together. In addition, although members of a social category may be different in some ways from members of other categories, these differences tend to become exaggerated and overgeneralized. Thus, categorization enhances perceptions of similarities within groups and differences between groups, emphasizing social difference and group distinctiveness. For social categorization, this process becomes more ominous because these within- and between-group distortions have a tendency to generalize to additional dimensions (e.g., character traits) beyond those that differentiated the categories originally (Allport, 1954). Furthermore, as the salience of the categorization increases, the magnitude of these distortions also increases (Brewer, 1979).

Upon social categorization of people as members of the ingroup and of outgroups, people favor ingroup members in reward allocations (Tajfel et al., 1971). Upon social categorization, people favor ingroup members, both explicitly and implicitly in evaluations (Otten and Moskowitz, 2000). Cognitively, people retain more information in a more detailed fashion for ingroup members than for outgroup members (Park and

Rothbart, 1982), have better memory for information about ways ingroup members are similar and outgroup members are dissimilar to the self (Wilder, 1981). In addition, people are more generous and forgiving in their explanations for the behaviors of ingroup relative to outgroup members. Positive behaviors and successful outcomes are more likely to be attributed to internal, stable characteristics (the personality) of ingroup than outgroup members, whereas negative outcomes are more likely to be ascribed to the personalities of outgroup members than of ingroup members (Pettigrew, 1979). Relatedly, observed behaviors of ingroup and outgroup members are encoded in memory at different levels of abstraction (Maass et al., 1989). Undesirable actions of outgroup members are encoded at more abstract levels that presume intentionality and dispositional origin (e.g., she is hostile) than identical behaviors of ingroup members (e.g., she slapped the girl). Desirable actions of outgroup members, however, are encoded at more concrete levels (e.g., she walked across the street holding the old man's hand) relative to the same behaviors of ingroup members (e.g., she is helpful).

In terms of social relations and behavioral outcomes, people are more cooperative and trustful of ingroup than outgroup members (Voci, 2006), and they exercise more [p. 445 ↓] personal restraint when using endangered resources shared with ingroup members than with others (Kramer and Brewer, 1984). Additionally, shared group membership decreases psychological distance from and facilitates arousal of promotive tension or empathy toward ingroup members (Hornstein, 1976). Relatedly, prosocial behavior is offered more readily to ingroup than to outgroup members (Dovidio et al., 1997; Piliavin et al., 1981).

Social categorization is a dynamic process, however, and people possess many different group identities and are capable of focusing on different social categories. By modifying a perceiver's goals, perceptions of past experiences, and/or expectations, there is opportunity to alter the level of category inclusiveness that will be primary or most influential in a given situation. This malleability of the level at which impressions are formed is important because of its implications for altering the way people think about members of ingroups and outgroups, and consequently about the nature of intergroup relations. Attempts to combat these biases can therefore be directed at altering the nature of social categorization.

The Common Ingroup Identity: The Theory and Initial Empirical Evidence

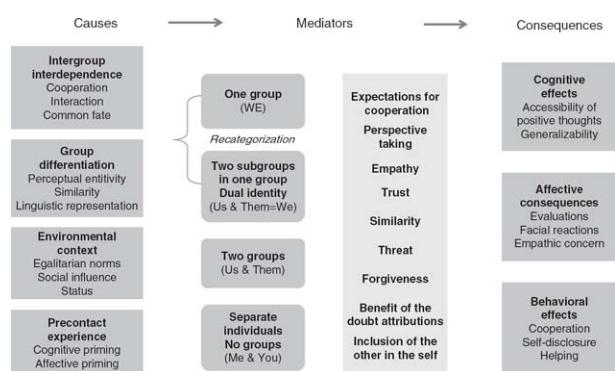
The key idea of the common ingroup identity model is that factors that induce members of different groups to recategorize themselves as members of the same more inclusive group can reduce intergroup bias through cognitive and motivational processes involving ingroup favoritism (Gaertner and Dovidio, 2000; Gaertner et al., 1993). Thus, more positive beliefs, feelings, and behaviors, which are usually reserved for ingroup members, are extended or redirected to former outgroup members because of their recategorized ingroup status. Consequently, recategorization dynamically changes the conceptual representations of the different groups from an “us” versus “them” orientation to a more inclusive, superordinate connection: “we.” Allport (1954) also recognized the potential value of perceiving inclusive ingroup membership across group lines when he asked hopefully, “Can a loyalty to [hu]mankind be fashioned before interracial warfare breaks out?”

The common ingroup identity model proposes (see [Figure 48.1](#)) that the different types of intergroup interdependence and cognitive, perceptual, linguistic, affective, and environmental factors, which include features specified by Allport's (1954) contact hypothesis (i.e., cooperative intergroup interaction, equal status, egalitarian norms), can either independently or in concert alter individuals' cognitive representations of the aggregate. In addition, common ingroup identity may be achieved by increasing the salience of existing common superordinate memberships (e.g., a team, a school, a company, a nation) or categories (e.g., students; Gómez et al., 2008) or by introducing factors (e.g., common goals or fate; see Gaertner et al., 1999) that are perceived to be shared by the memberships. These cognitive representations as one group, two subgroups within a more inclusive group (i.e., a dual identity), two separate groups, or separate individuals, are proposed to then produce specific cognitive (e.g., accessibility of positive thoughts), affective (e.g., evaluations) and overt behavioral consequences (e.g., self-disclosure and helping)

Once people regard former outgroup members as ingroup members, they are proposed initially to accord them benefits of ingroup status heuristically that, in turn, mediate

the relation between a one-group (or dual identity) representation and the ultimate cognitive, affective, and overt behavioral consequences. The benefits awarded heuristically to these new ingroup members include: decreased threat, increased empathy, trust, forgiveness, similarity to the self, increased inclusion of the other in the self, increased willingness to take the other's perspective, and more generous attributional interpretations [p. 446 ↓] for the other's positive and negative behavior. These benefits are conceived to be the mediating psychological processes that are engaged by the perception of common ingroup identity which ultimately result in more positive cognitive, affective, and behavioral orientations toward these former outgroup members.

Figure 48.1 The common ingroup identity model



We acknowledge that there may be other ways to address the role of social categorization to reduce intergroup bias. In their mutual intergroup differentiation model, Brown and Hewstone (2005) posit that intergroup relations will be harmonious when group identities remain salient but within a context of cooperative intergroup interaction. Alternatively, Wilder (1981) proposed that an effective way to reduce intergroup bias is through decategorization, in which the salience of group boundaries is weakened and people are encouraged to regard one another primarily as distinct individuals. Brewer and Miller (1984; see also Miller, 2002) emphasize the additional value of personalization, in which information about each other's unique qualities is exchanged for reducing bias. We see these approaches as alternative strategies which over time can often operate in complementary ways with the development of common identity.

One of our first experiments in this area directly explored how both common identity and deategorization can both reduce intergroup bias but in different ways (Gaertner et al., 1989). We compared the consequences of inducing two three-person *ad hoc* laboratory groups of college students to regard themselves as one group, two groups, or separate individuals. To manipulate these representations, we systematically varied a number of elements of the contact situation, including the spatial arrangement of the members (i.e., integrated, segregated or separated seating patterns) and the nature of the interdependence among the participants.

As we predicted, participants in the one-group and separate-individuals conditions reported lower bias (in liking and other evaluative characteristics) of the original ingroup and outgroup members relative to those in the two-groups condition. In addition, and as we hypothesized, participants in [p. 447 ↓] the one-group and separate-individuals conditions reduced bias in different ways. In the one-group condition, bias was reduced primarily because evaluations of former out-group members became more positive; in the separate-individuals condition, evaluations of former ingroup members became less positive. More recently, employing procedures very similar to those used in Gaertner et al. (1989), Guerra et al. (2010), and Rebelo et al. (2004) obtained very similar patterns of findings among 9 and 10-year-old Black and White groups of children in Portugal. The nearly identical patterns of findings across these separate studies conducted with *ad hoc* and racial groups, in different countries, at different times, among people of different age groups, and using different measures illustrate the generalizability of the effects we initially observed.

Encouraged by the findings of our 1989 experiment, which provided initial support for the common ingroup identity model by revealing how recategorization can reduce intergroup bias by increasing the attractiveness of former outgroup members, we thought about what features, beyond the walls of our laboratory, might induce more inclusive group representations among the members of different groups, especially different racial groups. For some time, psychologists have known that the conditions outlined in the contact hypothesis (Allport, 1954; see also Pettigrew, 1998; and Pettigrew and Tropp, 2006), involving cooperative intergroup interaction, equal status between the groups, opportunities for self-revealing interactions, and equalitarian norms endorsed by relevant authority enhance the effect of intergroup contact on reducing intergroup bias. What has been more elusive, however, has been the identification

of the psychological processes that are activated by these conditions of contact that mediate their relation to positive intergroup consequences.

It seemed reasonable to us that each of the conditions of contact specified by Allport (1954), which includes the cooperative intergroup interaction component specified in Sherif et al.'s functional model, may share the capacity to induce more inclusive cognitive representations among the memberships. In turn, these more inclusive representations are proposed to activate the series of psychological processes we specified above. This insight thereby extends Allport's contact hypothesis and Sherif's functional theory by linking their hypotheses to measurable psychological processes.

We directly tested these ideas with regard to the importance of intergroup cooperation in an experiment (Gaertner et al., 1990). This study again brought two three-person laboratory groups together under conditions designed to vary independently the members' representations of the aggregate as one group or two groups (by varying factors such as seating arrangement) and the presence or absence of intergroup cooperative interaction. Supportive of the hypothesis concerning how cooperation reduces bias, among participants induced to feel like two groups, the introduction of cooperative interaction increased their perceptions of one group and also reduced their bias in evaluative ratings, primarily by enhancing evaluations of outgroup members relative to those who did not cooperate during the contact period. Although Sherif et al. (1961) revealed that competition between groups increases intergroup conflict while intergroup cooperation produces intergroup harmony, the current study reveals just how cooperative intergroup interaction, in part, psychologically produces intergroup harmony by influencing the salience of intergroup boundaries. Supportive of the common ingroup identity model, these findings reveal that cooperation between groups changes the members' categorization scheme of the memberships from "us and them" to a more inclusive "we." Although Sherif and Sherif (1969: 286–288) recognized the capacity of intergroup cooperation to facilitate the development of a common superordinate entity, they conceived of this entity more concretely as the development of an emergent, formal organizational system with shared norms and standards, rather than a common identity that exists more ephemerally only in the mind of each member.

[p. 448 ↓]

The early evidence on the common ingroup identity model was generally strong in support for the model. However, the model was not without critics. For example, with respect to external validity, Hewstone (1996: 351) questioned whether a common ingroup identity can override powerful racial or ethnic categories on “more than a temporary basis.” Indeed, our initial support for the model was based primarily on research with groups formed in the laboratory. In the next section we address the issue of the effectiveness of common identity in improving relations between meaningful, enduring groups, and consider the model's potential for addressing social problems.

Addressing Social Problems

In this section, we examine research testing the applicability of the common ingroup identity model to relationships between members of groups in meaningful relations in “real-world” contexts and consider related challenges to the framework.

The Application of the Model to Real Groups

Since the initial research on the common ingroup identity model, a considerable body of work has demonstrated the applicability of the model to relations between a range of different groups in meaningful naturalistic contexts. Three survey studies across different domains of intergroup life offered converging support for the idea that features specified by the contact hypothesis increase intergroup harmony, in part because they transform members' representations of the memberships from separate groups to one more inclusive group. Participants in these studies included students attending a multiethnic high school (Gaertner et al., 1996), banking executives who had experienced a corporate merger involving a wide variety of banks across the US (Bachman, 1993), and college students who are members of blended families (Banker and Gaertner, 1998). In general, across these studies the more favorable participants reported the conditions of contact between the groups (e.g., cooperation), the more the school (or company or family) felt like one group. Supportive of the model, the more it felt like one group, the lower the bias in affective reactions in the high school, the less

the intergroup anxiety among the banking executives, and the greater the amount of stepfamily harmony. Also, a longitudinal study of stepfamilies found evidence supportive of the direction of causality between the constructs proposed by our model across time (Banker, 2002). Thus, we have found support for the model across a variety of intergroup settings and methodological approaches.

We also obtained experimental evidence of the effects of creating common identity on interracial behavior. One study was a field experiment conducted at the University of Delaware football stadium prior to a game between the University of Delaware and Westchester State University (Nier et al., 2001, Study 2). In this experiment, Black and White students approached fans of the same sex from both universities just before the fans entered the stadium. These fans were asked if they would be willing to be interviewed about their food preferences. Our student interviewers systematically varied whether they were wearing a University of Delaware or Westchester State University hat. By selecting fans who wore similar clothing that identified their university affiliation, we systematically varied whether fans and our interviewers had common or different university identities in a context where we expected these identities to be particularly salient. Although we planned to over-sample Black fans, the sample was still too small to yield any informative findings.

Among White fans, however, sharing common university identity with the Black interviewers significantly increased their compliance (59 percent) relative to when they did not share common identity with the Black interviewer (36 percent). When the [p. 449 ↓] interviewers were White, however, there was no significant difference in their levels of compliance as a function of their university identity. They gained equivalent levels of compliance when they shared common university identity with the fan (44 percent) as when they appeared to be affiliated with the rival university (37 percent). These findings offer support for the idea that outgroup members will be treated more favorably when they are perceived to also share a more inclusive, common ingroup affiliation.

In other work, we explored whether a common ingroup identity can affect the fundamental way that Whites think about Blacks during interracial interactions. Within the aversive racism framework, a major motive of Whites in interracial situations is to *avoid wrongdoing*. Supportive of this view, we have found across a variety of different studies that Whites typically do not discriminate against Blacks in situations in which

norms for appropriate behaviors are clearly defined. Thus, Whites can, at least under some circumstances, successfully suppress negative beliefs, feelings and behavior toward Blacks when it is obvious that expressing such reactions reflects racial bias. Unfortunately, in view of work on stereotype suppression and rebound, it is possible that once this self-imposed suppression is relaxed, negative beliefs, feelings, and behaviors would be even more likely than if they were not suppressed initially.

The common ingroup identity model, because it focuses on redirecting the forces of ingroup favoritism, can potentially change the motivational orientation or intentions of aversive racists from trying to avoid wrongdoing to trying to *do what is right*. Some preliminary evidence from our laboratory suggests the potential promise of a common ingroup identity to alter motivation in just such a positive way (Gaertner and Dovidio, 2000).

In this experiment, White participants who were about to interact with a White or a Black confederate were either asked to try to avoid wrongdoing, instructed to try to behave correctly toward the other person, informed that they were part of the same team with their partner and competing against a team at a rival institution, or were given no instructions. The dependent measure of interest was the relative accessibility of negative thoughts, as assessed by changes in responses on a Stroop color-naming task after the interaction relative to responses on a baseline Stroop task administered before the interaction. A rebound effect would be reflected in greater accessibility (i.e., longer color-naming latencies) of negative relative to positive words on the post-test Stroop task.

We hypothesized that because the primary motivation of aversive racists in interracial interaction is to avoid wrongdoing and thus to suppress negative thoughts and feelings, participants explicitly instructed to avoid wrongdoing and those given no instructions would show relatively strong accessibility of negative thoughts after interacting with a Black confederate. In contrast, we expected participants instructed to behave correctly and those in the “same team” condition (who were hypothesized to adopt a positive orientation on their own) would escape such a rebound effect.

The results revealed that when the confederate was White, the experimental conditions did not affect the accessibility of negative thoughts from one another or from baseline.

When the confederate was Black, however, the increased accessibility of negative relative to positive characteristics (from the pretest to the post-test) in the avoid-wrongdoing and no-instructions conditions was significantly greater than in the do-right and same-team conditions, in which there was an increase in the accessibility of positive relative to negative thoughts. The pattern of these findings suggests that the development of a common ingroup identity can alter motivation in interracial situations from one of suppressing negative thoughts, feelings, and actions to one that is positive, more appetitive and prosocial – and in a way that does not ironically result in further increases in negative thoughts. These findings are particularly encouraging to us because they illustrate the potential of a common ingroup identity for addressing the underlying [p. 450 ↓] motivational dynamics of aversive racism, which is where our journey began. Other researchers, however, also have recognized the potential of a common ingroup identity to address a variety of social issues.

Recategorization in terms of a common ingroup identity can promote intergroup forgiveness and trust. For instance, Wohl and Branscombe (2005) showed that increasing the salience of Jewish students' "human identity," in contrast to their "Jewish identity," increased their perceptions of similarity between Jews and Germans, as well as their willingness to forgive Germans for the Holocaust (Study 3) and their willingness to associate with contemporary German students (Study 4). A shared superordinate identity has also been shown to affect responsiveness to others. Kane et al. (2005) found that group members were more accepting of a newcomer's innovation when the newcomer shared a superordinate identity with them than when the newcomer did not, and that the strength of superordinate group identification was positively related to the extent to which group members accepted the innovative solution. Also, people are more responsive to the needs of former outgroup members perceived within a common ingroup identity (Dovidio et al., 1997) across a range of situations, including emergency situations (Levine et al., 2005).

In general, when we present this work people frequently question whether the development of a common ingroup identity is a realistic strategy to change and sustain more tolerant norms and attitudes. Our evaluation study of an elementary school antibias education intervention attempts to address this question.

Houlette et al. (2004) evaluated hypotheses derived from the common ingroup identity model in a quasi field experiment in the context of the Green Circle school-based antibias intervention program, which is designed to combat a range of biases (based on weight and sex, as well as race and ethnicity) among first- and second-grade children. The guiding assumption of the Green Circle Program, which is practically and theoretically compatible with the common ingroup identity model, is that helping children bring people from different groups conceptually into their own circle of caring and sharing fosters appreciation of their common humanity as well as respect for their differences. In particular, facilitators engage children in a variety of exercises designed to expand the circle, for instance, emphasizing, "All of us belong to one family – the human family." In terms of outcomes, the Green Circle intervention motivated the children to be more inclusive in selecting their most preferred playmate. Specifically, compared with children in the control condition who did not participate in Green Circle activities, those who were part of Green Circle showed significantly greater change in willingness to select other children who were different than themselves in race and in sex as a child that they "would most want to play with." These changes in the most preferred playmate involve a child's greater willingness to cross group boundaries in making friends – a factor that is one of the most potent influences in producing more positive attitudes toward the outgroup as a whole (Pettigrew, 1998).

In terms of future research applying our model to social problems, although racial disparity in health outcomes is a multifaceted problem, we are excited about the opportunity to develop an intervention based upon the common ingroup identity model to provide patients and physicians with a common team identity that has the potential to influence the communication between physician and patient as well as patient trust and adherence to physician recommendations (Penner et al., 2009).

Problems and Challenges

Although there is substantial evidence supportive of the usefulness of the common ingroup identity model for reducing intergroup bias and conflict, a number of problems and challenges remain to be addressed.

[p. 451 ↓]

Conflict at the Superordinate Level

It is important to recognize that the successful induction of a common ingroup identity does not necessarily eliminate social biases entirely. Upon recategorization, other outgroups, probably at the same level of inclusiveness, are likely to become the focus of “we–they” comparisons intended to maintain or protect the recategorized group’s positive distinctiveness. Kessler and Mummendey (2001), for example, found that East Germans who recategorized West Germans and East Germans as Germans became more biased over time toward members of other countries relative to those who continued to use the East–West German categorization scheme. Kessler and Mummendey noted that “recategorization is a 2-edged process: Although it reduces conflict at the subgroup level, it may initiate conflict at the common ingroup level” (2001: 1099). Thus, although recategorization may only redirect bias rather than reduce it completely, this may be desirable when conflict is especially unproductive.

Complementary and Reciprocal Processes

Although research on the model has generally focused on the immediate impact of creating a common ingroup identity, recategorization may operate in a complementary fashion with other processes over time to reduce intergroup bias in a more general and sustained way (Gaertner and Dovidio, 2000; Hewstone, 1996; Pettigrew, 1998). The more favorable impressions of and orientations toward outgroup members produced by recategorization within a common ingroup identity are not likely to be finely differentiated, at least initially. Rather, these more elaborated, personalized impressions can soon develop within the context of a common identity because the newly formed positivity bias is likely to encourage more open communication and greater self-disclosing interaction (see Dovidio et al., 1997) and friendship (see West et al., 2009) between former outgroup members. Thus, over time a common identity is proposed to encourage personalization of outgroup members (that reduces bias through a deategorization strategy that de-emphasizes group identities altogether; see Miller et al., 1985) and thereby initiates a second route to achieving reduced bias.

Intergroup Threat among Subgroups

In addition, efforts to induce a common identity can sometimes be met with resistance that can increase bias between members of the original groups. Social identity theory (Tajfel and Turner, 1979) proposes that people are motivated to maintain the positive distinctiveness of their group relative to other groups. When the integrity of one's group identity is threatened, people are motivated to re-establish positive and distinctive group identities and thereby maintain relatively high levels of intergroup bias (Brown and Wade, 1987) or show increased levels of bias (see Jetten et al., 2004 for a review). Consistent with this reasoning, introducing interventions such as emphasizing similarity or overlapping boundaries between the groups (Dovidio et al., 1997) or shared identity (Hornsey and Hogg, 2000) can exacerbate intergroup bias as a way of reaffirming positive distinctiveness. This effect is particularly likely to occur among people who value their original group highly, such as those more highly identified with their original group (Crisp et al., 2006), and when the initiative to form a superordinate identity is perceived to come from an outgroup rather than an ingroup member (Gómez et al., 2008). Additional work, however, suggests that once induced successfully, a common identity can reduce threat among subgroups. Riek et al. (2010) directly manipulated the salience of Democrats' and Republicans' shared identity as Americans and obtained experimental evidence suggesting that common identity increases positive outgroup attitudes by first reducing intergroup threat.

A Dual Identity

We note that the development of a common ingroup identity does not necessarily require each group to forsake its less inclusive group identity completely (Gaertner et al., 1989). [p. 452 ↓] Social identities are complex, and every individual belongs to multiple groups simultaneously (Brewer, 2000). Thus, depending on their degree of identification with different categories and contextual factors that make particular identities more salient, individuals may activate one or more of these identities simultaneously (Roccas and Brewer, 2002), as well as sequentially (Turner, 1985). As reflected in the "subgroups within one group" (i.e., a dual identity) representation,

it is possible for members to conceive of two groups (e.g., science and art majors) as distinct units (thereby establishing “mutual group differentiation”) within the context of a superordinate (i.e., university identity) social entity.

When group identities and their associated cultural values are adaptive, or when they are associated with high status or highly visible cues to group membership, it would be undesirable or impossible for people to relinquish these group identities completely or, as perceivers, to be “colorblind.” Indeed, demands to forsake these group identities or to adopt a colorblind ideology would likely arouse strong reactance and exacerbate intergroup bias. If, however, people continued to regard themselves as members of different groups, but all playing on the same team, intergroup relations between these “subgroups” would usually become more positive than if members only considered themselves as “separate groups.” Although, Allport (1954) observed that “concentric loyalties take time to develop, and often of course they fail completely to do so” (1954: 44–45), when subgroup identities are recognized, valued, and linked positively to the superordinate group identity, a dual identity may be effective for reducing intergroup bias and maintaining harmonious relations between groups. In addition, while it is not likely that a common ingroup identity can replace national, religious, and ethnic bases of categorization for very long, it is encouraging that its effects can be longer lasting, in part because of the reciprocal relation between the different categorization-based strategies (e.g., recategorization and personalization) that can continue to improve relations between the subgroups.

Conceptually, whereas the mutual intergroup differentiation model emphasizes the value of maintaining separate group identities within positive functional relations (i.e., cooperation) between groups, the common ingroup identity model posits that the superordinate identity component of a dual identity can be established in other ways, such as increasing the salience of overarching entities, as well. From the perspective of the common ingroup identity model it is the simultaneous salience of separate and superordinate group identities, not the particular mechanism that achieves this, that is important for intergroup bias. Consistent with this position, there is evidence that the intergroup benefits of a strong superordinate identity can be achieved for both majority and minority group members when the strength of the subordinate identity is also high (Huo et al., 1996). These findings are also conceptually consistent with studies that reveal that interethnic attitudes are more favorable when participants are primed with

a multicultural, pluralistic ideology for increasing interethnic harmony that emphasizes the value of a dual identity, relative to an assimilation ideology, which closely parallels a one-group representation (Richeson and Nussbaum, 2004).

Moreover, the benefits of intergroup contact may more easily generalize to additional outgroup members with a dual identity than with the pure one-group representation because the associative link to their original group identity remains intact, as in the mutual intergroup differentiation model. Research by González and Brown (2003, 2006) offers empirical support for this possibility. They found that manipulations designed to emphasize one group or dual identity representations produced equivalently low levels of bias in reward allocations to ingroup and outgroup members who were present in the contact situation. The researchers further investigated generalized bias, involving reward allocations to ingroup and outgroup members who were viewed on videotape. In terms of the relative [p. 453 ↓] amount of bias between conditions, generalized bias was somewhat, but not reliably, lower in the dual-identity condition than in the one-group condition. With respect to the extent to which participants within each condition favored their ingroup over the outgroup, there was no significant bias in the dual-identity condition, whereas bias was not fully eliminated in the one-group condition. Thus, a dual identity may offer some advantages over a one-group identity for generalized reductions in bias.

We note, however, that in contrast to the consistently positive relationship between the experience of a common identity (i.e., a one-group representation) and more favorable intergroup orientations, the strength of a dual identity can have divergent effects, associated with either positive or negative intergroup responses. For instance, in the multiethnic high-school study, a dual identity was related to lower bias, whereas in studies of banking executives involved in a merger and of members of blended families, a stronger sense of a dual identity was related to greater bias and conflict (see Gaertner et al., 2001).

One explanation for this latter effect is that when a common identity is made salient for members of different groups, members of one group or both groups may begin to regard their subgroup's characteristics (such as norms, values, and goals) as more prototypical of the common, inclusive category compared with those of the other subgroup (Mummendey and Wenzel, 1999). When this occurs, the outgroup is judged

as substandard, deviant, or inferior, leading to greater bias between the subgroups (e.g., Waldzus et al., 2004). This type of group projection that exacerbates bias may be more likely to occur when the superordinate identity represents a dimension directly relevant to the subcategory identities (e.g., Germans for East Germans and West Germans); when the superordinate identity is irrelevant to the subgroup identities, the experience of a dual identity is likely to have more favorable intergroup consequences (Hall and Crisp, 2005; see also Dovidio et al., 2008. Thus, further work is needed to understand the factors that can moderate the effectiveness of a dual identity for reducing bias and illuminate the mechanisms accounting for these effects.

Conclusion

In summary, the evidence reviewed in this chapter reveals consistent support for the key principle of the common ingroup identity model: successfully inducing ingroup and outgroup members to adopt a one-group representation inclusive of both groups, even when earlier group boundaries remain salient, reduces intergroup bias. Furthermore, this fundamental premise has been supported across studies using a variety of methodological approaches with participants of different ages, races and nationalities.

Nevertheless many questions, of both theoretical and practical significance, remain. For example, if recategorization as one group generalizes beyond those participants present (González and Brown, 2003, 2006), what is the mechanism through which this occurs? One explanation is that even when common identity is primary, vestiges of previous different group identities still remain. Thus, functionally, the major difference between the one-group and dual-identity representations may be one of degree – likely relating to the *relative* salience of both the subgroup and the superordinate identities.

Also, it would be important to learn whether it would be beneficial to activate recategorization processes when groups are engaged in direct, potentially mortal, conflict, and if so, how to effectively activate such a representation. Practically, Kelman (1999), whose work has focused on improving Palestinian–Israeli relations to achieve peace in the Middle East, has demonstrated that it is not necessary to create a common ingroup identity among *all* of the people involved in conflict to improve intergroup relations significantly. Kelman's conflict resolution workshops (see Rouhana and

Kelman, 1994) [p. 454 ↓] bring together 8–16 influential leaders from both sides in interactive, problem-solving exercises. These leaders can potentially create coalitions of peace-minded participants across conflict lines. And the workshops present a model for a new relationship between the parties (1994: 665). Thus even within the context of intense historical and contemporary conflict, it may be possible to be creative and to engineer the development of a common ingroup identity for a subset of group members with significant residual effects for the groups as a whole.

As the Hippocratic Oath proscribes, “Thou shalt do no harm,” we need to further identify negative consequences associated with recognition of common identity across group lines as well as factors that might mitigate these effects. For example, Wright (2001) proposed that recognition of common identity may mitigate a lower status group's desire for collective action to address social inequalities. Likewise, Saguy et al. (2009) suggest that commonality-focused contact among disadvantaged groups produced unrealistic expectations of fair and egalitarian treatment which were not forthcoming by members of the advantaged group.

In conclusion, this chapter describes the origins, development, applications, and challenges of the common ingroup identity model from both a theoretical and personal perspective. It reveals the importance of serendipity in scientific research – how unexpected findings can lead to new theoretical opportunities – and shows how different lines of research and professional interests can converge in unforeseeable ways. The chapter demonstrates the metamorphosis of our research interests from identifying a problem, aversive racism, to addressing the issue, the common ingroup identity model. The broader and more fundamental message that our work conveys, however, is that it is not only empirical support that advances theory but also the practical and conceptual challenges to the theory that stimulate conceptual development and new perspectives on age-old problems.

Acknowledgments

Preparation of this chapter was supported by NSF Grant # BCS–0613218.

Samuel L.Gaetner John F.Dovidio

References

- Adorno, T.W., Frenkel-Brunswik, E., Levinson, D.J. and Sanford, R.N. (1950) *The Authoritarian Personality*. New York: Harper.
- Allport, G.W. (1954) *The Nature of Prejudice*. Cambridge, MA: Addison-Wesley.
- Bachman, B.A. (1993) An intergroup model of organizational mergers . Unpublished PhD dissertation, Department of Psychology, University of Delaware, Newark.
- Banker, B.S. (2002) Intergroup conflict and bias reduction in stepfamilies: A longitudinal examination of intergroup relations processes . Unpublished PhD dissertation, Department of Psychology, University of Delaware, Newark.
- Banker, B.S. and Gaertner, S.L. Achieving step-family harmony: An intergroup relations approach . *Journal of Family Psychology* , 12, 310–325. (1998) <http://dx.doi.org/10.1037/0893-3200.12.3.310>
- Brewer, M.B. Ingroup bias in the minimal intergroup situation: A cognitive-motivational analysis . *Psychological Bulletin* , 86, 307–324. (1979) <http://dx.doi.org/10.1037/0033-2909.86.2.307>
- Brewer, M.B. (2000) Reducing prejudice through cross-categorization: Effects of multiple social identities . In S. Oskamp (ed.), *Reducing Prejudice and Discrimination: The Claremont Symposium on Applied Social Psychology* , pp. 165–183. Mahwah, NJ: Erlbaum.
- Brewer, M.B. and Miller, N. (1984) Beyond the contact hypothesis: Theoretical perspectives on desegregation . In N. Miller, ed. and M.B. Brewer (eds), *Groups in Contact: The Psychology of Desegregation* , pp. 281–302. Orlando: Academic Press.
- Brown, R. and Hewstone, M. (2005) An integrative theory of intergroup contact . In M.P. Zanna (ed.), *Advances in Experimental Social Psychology* , 37, 255–343. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2805%2937005-5>

Brown, R.J. and Wade, G. Superordinate goals and intergroup behavior: The effect of role ambiguity and status on intergroup attitudes and task performance . European Journal of Social Psychology , 17, 131–142. (1987) <http://dx.doi.org/10.1002/ejsp.2420170202>

Campbell, D.T. Common fate, similarity and other indices of the status of aggregates of persons as social entities . Behavioral Science , 3, 14–25. (1958) <http://dx.doi.org/10.1002/bs.3830030103>

Crisp, R.J., Walsh, J. and Hewstone, M. Crossed categorization in common ingroup contexts . Personality and Social Psychology Bulletin , 32, 1204–1218. (2006) <http://dx.doi.org/10.1177/0146167206289409>

Dollard, J., Doob, L., Miller, N., Mowrer, O.H. and Sears, R.R. (1939) Frustration and Aggression . New Haven, CT: Yale University Press. <http://dx.doi.org/10.1037/10022-000>

Dovidio, J.F. and Gaertner, S.L. (2004) Aversive racism . In M.P. Zanna (ed.), Advances in Experimental Social Psychology , 36, 1–51. San Diego: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2804%2936001-6>

Dovidio, J.F., Gaertner, S.L. and Saguy, T. Another view of “we”: Majority and minority group perspectives on a common ingroup identity . European Review of Social Psychology , 18, 296–330. (2008) <http://dx.doi.org/10.1080/10463280701726132>

Dovidio, J.F., Gaertner, S.L., Validzic, A., Matoka, K., Johnson, B. and Frazier, S. Extending the benefits of re-categorization: Evaluations, self-disclosure and helping . Journal of Experimental Social Psychology , 33, 401–420. (1997) <http://dx.doi.org/10.1006/jesp.1997.1327>

Dovidio, J.F. and Morris, W.N. Effects of stress and commonality of fate on helping behavior . Journal of Personality and Social Psychology , 31, 145–149. (1975) <http://dx.doi.org/10.1037/h0076236>

Fiske, S.T. and Taylor, S.E. (1991) Social Cognition , 2nd Edition. New York: McGraw-Hill.

Gaertner, S.L. Helping behavior and racial discrimination among liberals and conservatives . *Journal of Personality and Social Psychology* , 25, 335–341. (1973) <http://dx.doi.org/10.1037/h0034221>

Gaertner, S.L., Bachman, B.A., Dovidio, J.D. and Bunker, B.S. (2001). Corporate mergers and stepfamily marriages: Identity, harmony, and commitment . In M.A. Hogg, ed. and D. Terry (eds), *Social Identity in Organizations* , pp. 265–288. Oxford: Blackwell.

Gaertner, S.L. and Dovidio, J.F. The subtlety of white racism, arousal, and helping behavior . *Journal of Personality and Social Psychology* , 35, 691–707. (1977) <http://dx.doi.org/10.1037/0022-3514.35.10.691>

Gaertner, S.L. and Dovidio, J.F. (1986) The aversive form of racism . In J.F. Dovidio, ed. and S.L. Gaertner (eds), *Prejudice, Discrimination, and Racism* , pp. 61–89. Orlando: Academic Press.

Gaertner, S.L. and Dovidio, J.F. (2000) Reducing Inter-group Bias: The Common Ingroup Identity Model . Philadelphia: The Psychology Press.

Gaertner, S.L., Dovidio, J.F., Anastasio, P.A., Bachman, B.A. and Rust, M.C. (1993) The common ingroup identity model: Recategorization and the reduction of intergroup bias . In W. Stroebe, ed. and M. Hewstone (eds), *European Review of Social Psychology* , 4, 1–26. New York: John Wiley and Sons. <http://dx.doi.org/10.1080/14792779343000004>

Gaertner, S.L., Dovidio, J.F., Bunker, B.S., Rust, M.C., Nier, J.A. and Ward, C.M. (1997) Does pro-whiteness necessarily mean anti-blackness? In M. Fine, ed. , L. Powell, ed. , L. Weis, ed. , and M. Wong (eds), *Off White* , pp. 167–178. New York: Routledge.

Gaertner, S.L., Dovidio, J.F., Rust, M.C., Nier, J., Bunker, B., Ward, C.M., Mottola, G.R. and Houlette, M. Reducing intergroup bias: Elements of inter-group cooperation . *Journal of Personality and Social Psychology* , 76, 388–402. (1999) <http://dx.doi.org/10.1037/0022-3514.76.3.388>

Gaertner, S.L., Mann, J.A., Dovidio, J.F., Murrell, A.J. and Pomare, M. How does cooperation reduce intergroup bias? *Journal of Personality and Social Psychology*, 59, 692–704. (1990) <http://dx.doi.org/10.1037/0022-3514.59.4.692>

Gaertner, S.L., Mann, J.A., Murrell, A.J. and Dovidio, J.F. Reduction of intergroup bias: the benefits of recategorization . *Journal of Personality and Social Psychology*, 57, 239–249. (1989) <http://dx.doi.org/10.1037/0022-3514.57.2.239>

Gaertner, S.L., Rust, M.C., Dovidio, J.F., Bachman, B.A. and Anastasio, P.A. (1996) The Contact Hypothesis: The role of a common ingroup identity on reducing intergroup bias among majority and minority group members . In J.L. Nye, ed. and A.M. Brower (eds), *What's Social About Social Cognition?* pp. 230–360. Newbury Park, CA: Sage.

Gómez, A., Dovidio, J.F., Huici, C., Gaertner, S.L. and Cuardrado, I. The other side of We: When outgroup members express common identity . *Personality and Social Psychology Bulletin*, 34, 1613–1636. (2008)

González, R. and Brown, R. Generalization of positive attitude as a function of subgroup and superordinate group identification in intergroup contact . *European Journal of Social Psychology*, 33, 195–214. (2003)

González, R. and Brown, R. Dual identities and intergroup contact: Group status and size moderate the generalization of positive attitude change . *Journal of Experimental Social Psychology*, 42, 753–767. (2006)

Guerra, R., Rebelo, M., Monteiro, M.B., Riek, B.M., Maia, E.W., Gaertner, S.L. and Dovidio, J.F. How should intergroup contact be structured to reduce bias among majority and minority group children? *Group Processes and Intergroup Relations*, 13, 445–460. (2010) <http://dx.doi.org/10.1177/1368430209355651>

Hall, N.R. and Crisp, R.J. Considering multiple criteria for social categorization can reduce intergroup bias . *Personality and Social Psychology Bulletin*, 31, 1435–1444. (2005) <http://dx.doi.org/10.1177/0146167205276084>

Hewstone, M. (1996) Contact and categorization: Social psychological interventions to change intergroup relations . In C.N. Macrae, ed. , C. Stangor, ed. , and M. Hewstone (eds), *Stereotypes and Stereotyping* , pp. 323–368. New York: Guilford Press.

Hornsey, M.J. and Hogg, M.A. Subgroup relations: A comparison of mutual intergroup differentiation and common ungroup identity models of prejudice reduction . *Personality and Social Psychology Bulletin* , 26, 242–256. (2000) <http://dx.doi.org/10.1177/0146167200264010>

Hornstein, H.A. (1976) Cruelty and Kindness: A New Look at Aggression and Altruism . Englewood Cliffs, N.J.: Prentice-Hall.

Houlette, M., Gaertner, S.L., Johnson, K.M., Banker, B.S., Riek, B.M. and Dovidio, J.F. Developing a more inclusive social identity: An elementary school intervention . *Journal of Social Issues* , 60, 35–56. (2004) <http://dx.doi.org/10.1111/j.0022-4537.2004.00098.x>

Huo, Y.J., Smith, H.J., Tyler, T.R. and Lind, E.A. Superordinate identification, subgroup identification, and justice concerns: Is separatism the problem; is assimilation the answer? *Psychological Science* , 7, 40–45. (1996) <http://dx.doi.org/10.1111/j.1467-9280.1996.tb00664.x>

Jetten, J., Spears, R. and Postmes, T. Intergroup distinctiveness and differentiation: A meta-analytic integration . *Journal of Personality and Social Psychology* , 86, 862–879. (2004) <http://dx.doi.org/10.1037/0022-3514.86.6.862>

Kane, A.A., Argote, L. and Levine, J.M. Knowledge transfer between groups via personnel rotation: Effects of social identity and knowledge quality . *Organizational Behavior and Human Decision Processes* , 96, 56–71. (2005) <http://dx.doi.org/10.1016/j.obhd.2004.09.002>

Kelman, H.C. The interdependence of Israeli and Palestinian national identities: The role of the other in existential conflicts . *Journal of Social Issues* , 55, 581–600. (1999) <http://dx.doi.org/10.1111/0022-4537.00134>

Kessler, T. and Mummendey, A. Is there any scapegoat around? Determinants of intergroup conflict at different categorization levels . *Journal of Personality and Social Psychology* , 81, 1090–1102. (2001) <http://dx.doi.org/10.1037/0022-3514.81.6.1090>

Kramer, R.M. and Brewer, M.B. Effects of group identity on resource utilization in a simulated commons dilemma . *Journal of Personality and Social Psychology* , 46, 1044–1057. (1984) <http://dx.doi.org/10.1037/0022-3514.46.5.1044>

Kovel, J. (1970) *White Racism: A Psychohistory* . New York: Pantheon.

LeVine, R.A. and Campbell, D.T. (1972) *Ethnocentrism: Theories of Conflict, Ethnic Attitudes and Group Behavior* . New York: John Wiley.

Levine, M., Prosser, A., Evans, D. and Reicher, S. Identity and emergency intervention: How social group membership and inclusiveness of group boundaries shape helping behavior . *Personality and Social Psychology Bulletin* , 31, 443–453. (2005) <http://dx.doi.org/10.1177/0146167204271651>

Maass, A., Salvi, D., Arcuri, L. and Semin, G.R. Language use in intergroup contexts: The linguistic intergroup bias . *Journal of Personality and Social Psychology* , 57, 981–993. (1989) <http://dx.doi.org/10.1037/0022-3514.57.6.981>

Miller, N. Personalization and the promise of Contact Theory . *Journal of Social Issues* , 58, 387–410. (2002) <http://dx.doi.org/10.1111/1540-4560.00267>

Miller, N., Brewer, M.B. and Edwards, K. Cooperative interaction in desegregated settings: A laboratory analog . *Journal of Social Issues* , 41 (3) 63–75. (1985) <http://dx.doi.org/10.1111/j.1540-4560.1985.tb01129.x>

Mummendey, A. and Wenzel, M. Social discrimination and tolerance in intergroup relations: Reactions to intergroup difference . *Personality and Social Psychology Review* , 3, 158–174. (1999) http://dx.doi.org/10.1207/s15327957pspr0302_4

Nier, J.A., Gaertner, S.L., Dovidio, J.F., Banker, B.S. and Ward, C.M. Changing interracial evaluations and behavior: The effects of a common group identity .

Group Processes and Intergroup Relations , 4, 299–316. (2001) <http://dx.doi.org/10.1177/1368430201004004001>

Otten, S. and Moskowitz, G.B. Evidence for implicit evaluative in-group bias: Affect-based spontaneous trait inference in a minimal group paradigm . Journal of Experimental Social Psychology , 36, 77–89. (2000) <http://dx.doi.org/10.1006/jesp.1999.1399>

Penner, L.A., Dailey, R.K., Markova, T., Porcerelli, H.H. Dovidio, J.F. and Gaertner, S.L. (2009) Using the common group identity model to increase trust and commonality in racially discordant medical interactions . Poster presented at the Annual Meeting of the Society for Personality and Social Psychology, Tampa, FL. (February 7).

Pettigrew, T.F. The Ultimate Attribution Error: Extending Allport's cognitive analysis of prejudice . Personality and Social Psychology Bulletin , 5, 461–476. (1979) <http://dx.doi.org/10.1177/014616727900500407>

Pettigrew, T.F. Intergroup contact theory . Annual Review of Psychology , 49, 65–85. (1998) <http://dx.doi.org/10.1146/annurev.psych.49.1.65>

Pettigrew, T.F. and Tropp, L. A meta-analytic test of intergroup contact theory . Journal of Personality and Social Psychology , 90, 751–783. (2006) <http://dx.doi.org/10.1037/0022-3514.90.5.751>

Piliavin, J.A., Dovidio, J.F., Gaertner, S.L. and Clark, R.D., III. (1981) Emergency Intervention . New York: Academic Press.

Piliavin, J.A. and Piliavin, I.M. (1973) The Good Samaritan: Why does he help? Unpublished manuscript, Department of Sociology, University of Wisconsin, Madison.

Rebelo, M., Guerra, R. and Monteiro, M.B. (2004) Reducing Prejudice: Comparative Effects of Three Theoretical Models . Paper presented at the 5th Biennial Convention of the Society for the Psychological Study of Social Issues, Washington, DC, (June 27).

Richeson, J.A. and R.J.Nussbaum The impact of multiculturalism versus color-blindness on racial bias . *Journal of Experimental Social Psychology* , 40, 417–423. (2004) <http://dx.doi.org/10.1016/j.jesp.2003.09.002>

Riek, B.M., Mania, E.W. and Gaertner, S.L. Intergroup threat and outgroup attitudes: A meta-analytic review . *Personality and Social Psychology Review* , 10, 336–353. (2006) http://dx.doi.org/10.1207/s15327957pspr1004_4

Riek, B.M., Mania, E.W., Gaertner, S.L., Direso, S.A. and Lamoreaux, M.J. Does a common ingroup identity reduce intergroup threat? *Group Processes and Intergroup Relations* , 13, 403–423. (2010) <http://dx.doi.org/10.1177/1368430209346701>

Roccas, S. and Brewer, M.B. Social identity complexity . *Personality and Social Psychology Review* , 6, 88–106. (2002) http://dx.doi.org/10.1207/S15327957PSPR0602_01

Rouhana, N.N. and Kelman, H.C. Promoting joint thinking in international conflicts: An Israeli-Palestinian continuing workshop . *Journal of Social Issues* , 50, 157–178. (1994) <http://dx.doi.org/10.1111/j.1540-4560.1994.tb02403.x>

Saguy, T., Tausch, N., Dovidio, J.F. and Pratto, F. The irony of harmony: Intergroup contact can produce false expectations for equality . *Psychological Science* , 20, 114–121. (2009) <http://dx.doi.org/10.1111/j.1467-9280.2008.02261.x>

Sherif, M., Harvey, O.J., White, B.J., Hood, W.R. and Sherif, C.W. (1961) *Intergroup Conflict and Cooperation: The Robbers Cave Experiment* . Norman, OK: University of Oklahoma Book Exchange.

Sherif, M. and Sherif, C.W. (1969) *Social Psychology* . New York: Harper and Row.

Steiner, I.D. Whatever happened to the group in social psychology?
Journal of Experimental Social Psychology , 10, 94–108. (1974). <http://dx.doi.org/10.1016/0022-1031%2874%2990058-4>

Tajfel, H. Cognitive aspects of prejudice . *Journal of Social Issues* , 25, 79–97. (1969). <http://dx.doi.org/10.1111/j.1540-4560.1969.tb00620.x>

Tajfel, H. and Turner, J.C. (1979) An integrative theory of intergroup conflict . In W.G. Austin, ed. and S. Worchel (eds), *The Social Psychology of Intergroup Relations* , pp. 33–48. Monterey, CA: Brooks/Cole.

Tajfel, H., Billig, M.G., Bundy, R.F. and Flament, C. Social categorisation and intergroup behavior . *European Journal of Social Psychology* , 1, 149–177. (1971) <http://dx.doi.org/10.1002/ejsp.2420010202>

Turner, J.C. (1985) Social categorization and the self-concept: A social cognitive theory of group behavior . In E.J. Lawler (ed.), *Advances in Group Processes* , 2, 77–122. Greenwich, CT: JAI Press.

Voci, A. The link between identification and ingroup favouritism: Effects of threat to social identity and trust-related emotions . *British Journal of Social Psychology* , 45, 265–284. (2006) <http://dx.doi.org/10.1348/014466605X52245>

Waldzus, S., Mummeney, A., Wenzel, M. and Boettcher, F. Of bikers, teachers, and Germans: Groups' diverging views about their prototypicality . *British Journal of Social Psychology* , 43, 385–400. (2004) <http://dx.doi.org/10.1348/0144666042037944>

West, T.V., Pearson, A.R., Dovidio, J.F., Shelton, N.J. and Trail, T.E. Strength of superordinate identity and intergroup roommate friendship development . *Journl of Experimental Social Psychology* , 45, 1266–1272. (2009) <http://dx.doi.org/10.1016/j.jesp.2009.08.002>

Wilder, D.A. (1981) Perceiving persons as a group: Categorization and intergroup relations . In D.L. Hamilton (ed.), *Cognitive Processes in Stereotyping and Intergroup Behavior* , pp. 213–257. Hillsdale, NJ: Erlbaum.

Wohl, M.J.A. and Branscombe, N.R. Forgiveness and collective guilt assignment to historical perpetrator groups depend on level of social category inclusiveness . *Journal of Personality and Social Psychology* , 88, 288–303. (2005) <http://dx.doi.org/10.1037/0022-3514.88.2.288>

Wright, S.C. (2001) Collective action: Social psychology and social change . In R. Brown, ed. and, S.L. Gaertner (eds), Blackwell Handbook of Social Psychology, Vol. 4: Intergroup Processes , pp. 409–430. Oxford: Blackwell.

<http://dx.doi.org/10.4135/9781446249222.n48>

Handbook of Theories of Social Psychology

Social Role Theory

Contributors: Alice H. Eagly & Wendy Wood

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Social Role Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n49>

Print pages: 458-477

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n49>

[p. 458 ↓]

Chapter 49: Social Role Theory

Abstract

What causes sex differences and similarities in behavior? At the core of our account are societal stereotypes about gender. These stereotypes, or gender role beliefs, form as people observe male and female behavior and infer that the sexes possess corresponding dispositions. For example, in industrialized societies, women are more likely to fill caretaking roles in employment and at home. People make the correspondent inference that women are communal, caring individuals. The origins of men's and women's social roles lie primarily in humans' evolved physical sex differences, specifically men's size and strength and women's reproductive activities of gestating and nursing children, which interact with a society's circumstances and culture to make certain activities more efficiently performed by one sex or the other. People carry out gender roles as they enact specific social roles (e.g., parent, employee). Socialization facilitates these sex-typical role performances by enabling men and women to develop appropriate personality traits and skills. Additionally, gender roles influence behavior through a biosocial set of processes: hormonal fluctuations that regulate role performance, self-regulation to gender role standards, and social regulation to others' expectations about women and men. Biology thus works with psychology to facilitate role performance.

Social Role Theory of Sex Differences and Similarities

A profound question about human life is why men and women, and boys and girls, behave differently in many circumstances but similarly in others. There is no one discipline that provides a sovereign, overarching answer, but each discipline favors

certain types of causes. For biologists, sex differences reflect gonadal or other sex-differentiated hormones. For sociologists, the differences reflect the position of men and women in broader social hierarchies. For economists, the differences reflect the human capital of women and men. For developmental researchers, they arise from sex-linked temperament and socialization experiences. Evolutionary psychologists usually favor sex-differentiated selection pressures on human ancestors. Our theory begins from a uniquely social psychological vantage point that highlights social roles and interweaves role-related processes with these other perspectives to produce a powerful analysis of sex differences and similarities.

[p. 459 ↓]

In brief, we argue that sex differences and similarities in behavior reflect *gender role beliefs* that in turn represent people's perceptions of men's and women's *social roles* in the society in which they live. In postindustrial societies, for example, men are more likely than women to be employed, especially in authority positions, and women are more likely than men to fill caretaking roles at home as well as in employment settings. Men and women are differently distributed into social roles because of humans' evolved *physical sex differences* in which men are larger, faster, and have greater upper-body strength, and women gestate and nurse children. Given these physical differences, certain activities are more efficiently accomplished by one sex or the other, depending on a society's circumstances and culture. This task specialization produces an alliance between women and men as they engage in a division of labor. Although these alliances take somewhat different forms across cultures, task specialization furthers the interests of the community as a whole.

Gender role beliefs arise because people observe female and male behavior and infer that the sexes possess corresponding dispositions. Thus, men and women are thought to possess attributes that equip them for sex-typical roles. These attributes are evident in consensually-shared beliefs, or *gender stereotypes*. In daily life, people carry out these gender roles as they enact specific social roles such as parent or employee. Because gender roles seem to reflect innate attributes of the sexes, they appear natural and inevitable. With these beliefs, people construct gender roles that are responsive to cultural and environmental conditions yet appear, for individuals within a society, to be stable, inherent properties of men and women.

To equip men and women for their usual family and employment roles, societies undertake extensive *socialization* to promote personality traits and skills that facilitate role performance. Additionally, gender roles influence behavior through a trio of biological and psychological processes. Biological processes include *hormonal fluctuations* that act as chemical signals that regulate role performance. Psychological processes include individuals' internalization of gender roles as *self standards* against which they regulate their own behavior as well as their experience of other people's *expectations* that provide social regulatory mechanisms. Biology thus works with psychology to facilitate role performance.

The broad scope of our theory enables it to tackle the various causes of female and male behavior that are of interest across the human sciences. But the theory was not developed all in one piece. As we explain, Alice initially developed its core components in the 1980s, drawing largely on work in psychology and sociology. We have since worked together to elaborate the model so that it addresses causation at several levels of analysis. That is, we have placed the theory in a broader *nomological net*, or series of connected theoretical concepts and observable properties that give the constructs particular meaning (Cronbach and Meehl, 1955). By looking upward in the net toward the distal, fundamental causes of sex differences and similarities, we can answer big-picture questions about the evolutionary origins of male and female roles. We also can look downward to understand how men and women enact behavior through proximal psychological and biological processes.

In its scope, our analysis is broader than the more focused theories typical in social and personality psychology, which are suited to explain more fine-grained issues of cognition, affect, and social interaction. Our approach thus explains the ultimate origins of sex differences in behavior. It also shows how the position of women and men in the social structure determines the particular content of the cognitions (i.e., gender role beliefs) that influence female and male behavior. In addition, the theory identifies the psychological and biological processes that act as proximal determinants of sex differences and similarities.

[p. 460 ↓]

Personal Narrative and Intellectual Context

When Alice began work on the psychology of gender in the late 1970s, social psychologists had paid very little attention to this topic. This fundamental aspect of human life had gone virtually unanalyzed. On those few occasions when social psychologists separated their data according to the sex of their research participants, the differences that sometimes emerged were puzzling and awkward. From a contemporary perspective, this lack of attention to gender is an amazing blind spot in the subdiscipline of psychology most concerned with understanding social life. Yet, the important historical developments in the field were essentially gender blind. Gordon Allport (1954), in his influential book on prejudice, did not consider gender prejudice and instead focused his analytical powers on ethnic and racial prejudice. In Heider's (1958) classic book on social interaction, the terms *sex* and *gender* do not even appear in the index. As social psychologists turned to cognitive consistency theories in the 1960s and to attribution and social cognition in the 1970s, their work still did not broach the topic of gender.

In contrast with the lack of academic attention, the second-wave feminist movement in the 1960s and 1970s (e.g., Friedan, 1963) made gender a major theme of public discourse in the US. The field of social psychology could not remain on the sidelines of such an engaging conversation. Gender was emerging as a significant political, social, and psychological issue. Nonetheless, there were risks in initiating research in such a neglected area of study. Alice already had attained a tenured university position and established expertise in attitudes, one of social psychology's mainstream research areas. So, hedging her bets and joined by a few other soon-to-be-distinguished scholars (e.g., Bem, 1974; Deaux, 1976), she decided that the potential for contributing to an important set of social and scientific issues outweighed any career risk. In fact, her first major journal article on gender (Eagly, 1978) broke new ground, as confirmed by its winning of two prizes (1976 Gordon Allport Award of Society for the Psychological Study of Social Issues; 1978 Distinguished Publication Award of Association for Women in Psychology). This response was encouraging.

Psychology had produced some significant scholarship on sex and gender in earlier decades, primarily within mental skills testing (Hollingworth, 1916), developmental

psychology (Maccoby, 1966), and psychoanalysis (e.g., Horney, 1967). Yet, these perspectives did not incorporate the core message of social psychology – the power of the situation in its interactions with the attributes and processes of individuals. In this void, Alice developed social role theory, which was initially published in a book based on lectures that she gave at the University of Alberta in 1985 (Eagly, 1987). Wendy Wood was a graduate student with Alice during this period, allowing a productive collaboration on gender to begin.

Core Gender Role Theory

The role concept was crucial to Alice's initial thinking about gender. This reflected her educational background in Harvard and University of Michigan programs that integrated psychological and sociological traditions of social psychology. Alice had studied role theorists in sociology extending back to the writers such as Georg Simmel, George Herbert Mead, Ralph Linton, and Jacob Moreno (see Biddle, 1979) and including then-contemporary writers such as Erving Goffman (1959). As reflected in this tradition, role is a central integrative concept in the social sciences that is important because of the analytical bridge it provides between the individual and the social environment. Role expectations thus exist in the minds of individuals and also are shared with other people, producing the social consensus from which social structure and culture emerge. The role concept thus facilitated a theory of [p. 461 ↓] gender that analyzes not only the proximal determinants of male and female behavior but also the more distal influences of culture and social structure that contribute to variability in this behavior.

Within the traditions of role theory, Parsons and Bales (1955) had provided an explicit analysis of female and male roles. These theorists described the division of labor between husbands and wives as a specialization of men in task-oriented (or instrumental) behavior and of women in socioemotional (or expressive) behavior. Allied researchers observed that in mixed-sex groups, men, more than women, specialized in instrumental behaviors related to task accomplishment, and women, more than men, in socioemotional behaviors related to group maintenance and other distinctively social concerns (Strodtbeck and Mann, 1956). Reasoning that role differentiation along these lines is functionally necessary to harmonious social interaction, Parsons and his collaborators viewed these complementary male and female roles as inherent in a

smoothly functioning society. As an undergraduate student in one of Parsons' courses, Alice was struck by the power of this analysis, but she also came to realize that it was incomplete. What was missing is an appreciation of the malleability of role structures. Gender roles – that is, expectations for female and male behavior – are not stuck inevitably in the 1950s American form that Parsons and Bales observed. Instead, these expectations change, depending on the typical work and family roles of the sexes. Parsons and Bales had captured a moment in time in a particular cultural context. Change in the work and family roles of men and women follows from the exigencies of the economy, technology, and broader social structure in which these roles are embedded.

Another early catalyst of our theory of gender was a methodological innovation – the late 1970s development of quantitative methods for systematically integrating research findings (e.g., Glass et al., 1981; Rosenthal, 1978). Application of these methods allowed researchers to make more definitive statements concerning female and male behavior. Comparing women and men was an easy application of meta-analytic methods, requiring only a relatively simple two-group, between-subjects comparison. Alice was an early adopter of meta-analysis, initially producing an integration of studies that had compared the influenceability of women and men in conformity and persuasion paradigms (Eagly and Carli, 1981). After Wendy landed her first job at a university without a strong undergraduate research participant pool, a condition that limited her access to participants, she also became impressed with the power of meta-analysis. We each published a number of meta-analytic projects, including ones comparing women and men on aggressive behavior (Eagly and Steffen, 1986), helping behavior (Eagly and Crowley, 1986), group performance (Wood, 1987), and happiness and life satisfaction (Wood et al., 1989).

The sex differences documented in these early meta-analytic investigations required explanation. Although the differences typically were not large when averaged across studies, they were relatively large in some settings, with some interaction partners, and with some forms of the behavior under investigation. Even the average differences were often large enough to be consequential, in view of the substantial cumulative impact that small differences can have if repeatedly enacted over a period of time (Abelson, 1985). Despite being sympathetic to the prevailing view among many psychologists that most sex differences are small, we found the aggregated sex differences and their variability

across studies in our meta-analyses to be anything but trivial. Instead, such data posed puzzles to be solved with the aid of relevant theory.

Another input into the beginnings of social role theory was the emergence of psychological research on cultural stereotypes about women and men. Although such work began in the 1950s (McKee and Sherriffs, 1957), it intensified and gained visibility in the 1970s (e.g., Broverman et al., 1972; Spence and Helmreich, 1978). This research identified people's consensual beliefs concerning men [p. 462 ↓] and women. Most of these beliefs can be summarized in two dimensions, which are often labeled *agentic* and *communal* (Bakan, 1966). Men, more than women, are thought to be agentic – that is, masterful, assertive, competitive, and dominant. Women, more than men, are thought to be communal – that is, friendly, unselfish, concerned with others, and emotionally expressive. These qualities are similar to those that Parsons and Bales (1955) had labeled as instrumental and expressive (or task-oriented and socioemotional). As abstract, general beliefs about men and women, these stereotypes constitute gender roles.

Our emerging realization that these stereotypes are neither arbitrary nor essentially inaccurate was buttressed by social psychological research showing that social perceivers usually assume that people's behaviors reflect their intrinsic characteristics. This cognitive process of inferring traits from observed behavior is known as *correspondent inference* or *correspondence bias* (Gilbert and Malone, 1995; Ross, 1977). This process is widespread (Gawronski, 2003) and largely spontaneous (Uleman et al., 2008). For example, upon observing an act of kindness, perceivers automatically identify the behavior in trait terms and characterize the actor by the trait that is implied – as a nice, caring person.

Consistent with correspondent inference, we recognized that in various ways the new meta-analytic findings pertaining to male and female behavior matched gender stereotypic findings reflecting people's beliefs about men and women. The behavioral differences thus resembled the beliefs that people hold about differences (Eagly and Wood, 1991). Confirmation of this informal observation followed in research by Swim (1994) and Hall and Carter (1999) that found substantial correlations between participants' beliefs about sex differences and the differences established in meta-analytic reviews. This similarity between gender stereotypes and male and female

behavior challenged social psychologists' traditional depiction of stereotypes as inaccurate portrayals of groups (Allport, 1954).

Our understanding of how gender stereotypic beliefs can in turn guide behavior was aided by research in psychology and sociology. Psychological research outlined the power of expectancies to produce behavior that confirms them (e.g., Rosenthal and Rubin, 1978; see review by Olson et al., 1996). That is, stereotypes can act as self-fulfilling prophecies. Sociological research featured expectation states theory (Berger et al., 1980), which also linked beliefs about social groups (e.g., sex, race) to the behavior of individual group members. Cecilia Ridgeway, a sociologist working within this tradition and Wendy's colleague for a few years, demonstrated how beliefs about men's greater worth and value, which are based on men's greater access to societal resources and power, produce sex differences in influence in small task-performing groups (Ridgeway, 1981, 1984). These psychological and sociological ideas, although very different in form, provided frameworks to understand how cultural beliefs about gender guide individuals' behavior to yield confirmatory evidence of sex differences in the context of social interaction.

During the years in which we developed our theory, vivid experimental demonstrations of the power of gender roles cumulated in the research literature. For example, Zanna and Pack's (1975) experiment showed that female students shaped their self-presentations to fit the preferences of a highly eligible male interaction partner. When this man reported preferring women who were traditional (versus nontraditional), these young women presented themselves as conforming to his preferences and furthermore scored worse on a test of intellectual aptitude given that these scores were to be shared with this male partner. In an experiment by Skrypnek and Snyder (1982), task partners negotiated a more traditional division of labor when they believed that their (unseen) partner was of the other sex, regardless of their partner's actual sex.

Yet another catalyst of social role theory was research showing that people's self-concepts tend to have gender-stereotypic content (e.g., Bem, 1974; Spence and [p. 463 ↓] Helmreich, 1978). Researchers had turned to gender stereotypes to choose items for measuring instruments that assess the ascription of agentic and communal attributes to the self. It thus appeared that, to varying extents, people internalize gender roles as personal gender identities. Wendy later pursued these ideas, evaluating whether

gender identities could serve as personal standards for behavior. Her empirical studies showed that, as with other self-regulatory standards (e.g., Carver and Scheier, 1981), men and women regulate their own behavior to correspond to these identities (see Witt and Wood, 2010; Wood et al., 1997).

In summary, Alice built the core concepts of social role theory in the 1980s from a variety of theoretical perspectives and empirical traditions: sociological role theory; research on gender stereotypes; ideas about correspondent inference, behavioral confirmation, and status construction; studies of gender identity and self-regulatory processes; and the methodological innovation of meta-analysis. Indicating the success of this approach, it effectively explained variability in sex differences observed in meta-analyses in the 1980s (Eagly and Crowley, 1986; Eagly and Steffen, 1986; Wood, 1987; Wood et al., 1989). Also, a number of key experimental demonstrations in our own research programs showed the power of these principles to account for variability in perceived and actual sex differences (Eagly and Steffen, 1984; Eagly and Wood, 1982; Eagly et al., 1981; Grossman and Wood, 1993; Wood and Karten, 1986).

The invitation that Alice received to deliver a series of lectures in the fall of 1985 at the University of Alberta included the agreement that the lectures would be turned into a book. The result was a book presenting social role theory along with supportive evidence from meta-analyses and primary research (Eagly, 1987). This invitation was very helpful because the lectures and associated book required a systematic theoretical presentation. This book turned out to be influential as a substantial theoretical statement, as indicated by its 1075 citations recorded in Web of Science and 1906 in Google Scholar.

Biosocial Mechanisms

More recently, we enlarged the scope of social role theory to address the origins of the male/female division of labor. A key development in psychology that spurred this growth was the emergence of evolutionary psychology, which provided an essentialist explanation of many sex differences in social behavior (e.g., Buss and Schmitt, 1993). Rather than leave the question of ultimate origins to evolutionary psychology, Wendy and Alice collaborated in challenging its proponents by providing an alternative origin

theory that treated female and male behavior as emergent from interacting social and biological causes. This work expanded our theory by considering the distal, evolutionary causes of gender roles.

Our initial foray into the origins question was spurred by Wendy's discussions with her running partner, colleague, and evolutionary psychologist, Jeff Simpson. During their 8-mile Sunday morning runs – which, to Jeff's credit, rarely involved arguments – Wendy was inspired to take up the question of the evolutionary origins of gender roles. At that same time, Alice gave an invited address to the Midwestern Psychological Association that considered the origins of sex differences in social behavior. Eager for some empirical support for our developing ideas about cultural influences on mate preferences, Alice asked David Buss to share data from his wellknown 37 cultures study (Buss, 1989). Alice's 1997 talk featured an initial reanalysis of these data, a small foray that then was considerably enlarged and refined in our subsequent article (Eagly and Wood, 1999).

In reanalyzing the 37 cultures data, we found that in societies with a strong division of labor between male providers and female homemakers (i.e., less gender equality), women were more likely to prefer a mate with resources who could be a good provider, and men were more likely to prefer a mate who was a skilled homemaker and child caretaker. This marital system of a good provider paired with a domestic worker also generated [p. 464 ↓] a spousal age difference, given that older men were more likely to have acquired resources, and younger women without resources were more likely to value marriage and older partners with resources. This project thus demonstrated that sex differences in mate preferences, assumed by evolutionary psychologists to stem from sexual selection pressures acting on the human species, reflected the position of women and men in the social structure.

This initial foray into cross-cultural analyses merely whetted our appetite for developing a biosocial evolutionary theory that included as a central component the idea of variability across cultures in men's and women's roles. To expand our analysis and gain additional empirical support, we turned to the anthropological literature on cross-cultural uniformity and variability in female and male behavior. We found rich data and theorizing about the origins of sex differences, much of it compatible with our emerging

biosocial perspective. The resulting article presented social role theory as a core set of ideas within the larger biosocial theoretical framework (Wood and Eagly, 2002).

Alice took social role theory in yet another new direction by considering the behavior of men and women in organizational environments, where they act under the influence of specific occupational roles along with gender roles. Although analysis of male and behavior in complex natural settings offers many possibilities, Alice narrowed her focus to consider the conflux of gender roles and managerial roles. A key insight was that for men, managerial (or leader) roles and their own gender role are similar in content but for women, these roles are dissimilar. Female leaders' resulting role incongruity has varied consequences, including prejudice toward them as potential leaders and actual leaders (Eagly and Karau, 2002).

Our most recent collaboration further developed the biosocial roots of gender roles by elaborating the proximal biological and social processes that yield female and male behavior. Wendy spent a year at the Radcliffe Institute for Advanced Study and, listening to other Institute Fellows talk about their work, she was challenged by the variety of ways that human sciences treat the biology of sex and the sociality and psychology of gender. This experience inspired us to scrutinize the ways in which gender roles influence behavior. Given the growth of science pertaining to the hormonal regulation of female and male behavior (e.g., Archer, 2006; Hines, 2009; Taylor et al., 2000), any theory of the proximal determinants of sex differences that relied only on social psychological mechanisms is incomplete. Therefore, to enhance the biological side of our biosocial framework, we drew on this new science to document how hormones are recruited to facilitate the performance of social roles (Wood and Eagly, 2010).

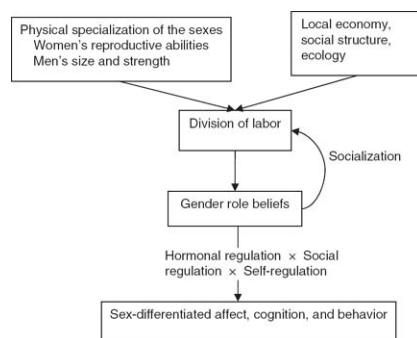
In summary, the 1980s version of social role theory (Eagly, 1987) has remained intact within a larger biosocial theoretical structure (see [Figure 49.1](#)). By reaching upward in the nomological net, we expanded the theory to include the evolutionary origins of the male/female division of labor. By reaching downward, we incorporated the emerging science on the hormonal regulation of social behavior. In addition, the analysis of leadership has provided a model of how the theory could take into account the interaction between gender roles and specific social roles in natural settings. In

the next section of this chapter, we offer a summary of the current theory and a brief acknowledgement of some of the relevant empirical literatures.

Social Role Theory of Sex Differences and Similarities: A Biosocial Approach

As indicated by the intellectual history presented in the prior section, our biosocial theory consists of a series of interconnected causes of sex differences and similarities. These causes range from more proximal (or immediate), to more distal (or ultimate). In Figure 49.1, the more distal causes are positioned above the [p. 465 ↓] division of labor, which is the outcome of interaction between the physical specialization of the sexes and local conditions. The division of labor yields gender role beliefs, which then facilitate this division through socialization processes. Gender role beliefs act on behavior through a trio of processes involving regulation by hormonal changes, others' expectations, and self standards. In this section, we begin our presentation of this theory with the ultimate determinants of female and male behavior and move to the more proximal determinants.

Figure 49.1 Gender roles guide sex differences and similarities through biosocial processes



Origins of the Division of Labor

The ultimate origins of male and female behavior derive from evolved physical differences between the sexes, especially women's reproductive activities and men's greater size and strength, as these factors interact with the demands of people's social and economic environment (Wood and Eagly, 2002). This interaction yields constraints whereby one sex performs certain tasks more efficiently than the other sex in a given environment.

Women's reproductive activities of pregnancy and lactation act as powerful constraints because they cede to women the energy-intensive and time-consuming activities of gestating, nursing, and caring for infants (Huber, 2007). These activities make it difficult for women to participate as fully as men in tasks that require speed of locomotion, uninterrupted activity, extended training, or long distance travel away from home. Therefore, in foraging, horticultural, and agricultural societies, women generally participate relatively little in tasks such as hunting large animals, conducting warfare, and plowing. Instead, women favor activities more compatible with childcare (see Murdock and Provost, 1973).

Also, men's greater size and strength equip them to execute tasks that benefit from these qualities, including hunting large animals, plowing, and warfare. Therefore, the division of labor between women and men reflects the specialization of each sex in activities for which they are physically better suited under the circumstances presented by their society. Because these circumstances are variable, the particular activities allocated [p. 466 ↓] to women and men differ across cultures (Wood and Eagly, 2002).

In addition to dividing tasks between women and men, societies often, but not always, cede greater power or status to men. In decentralized, nonhierarchical foraging or pastoral societies with limited technology, egalitarian relations between the sexes are common (Hayden et al., 1986; Salzman, 1999). Patriarchy arises when the physical attributes of men and women interact with economic and technological developments to give men the roles that yield decision-making authority (e.g., in warfare) and access to resources (e.g., through intensive agriculture and trade). Under such circumstances, men garner most of the social and economic capital that derives from these activities.

In recent history, both the division of labor and gender hierarchy, especially in industrialized societies, have become weaker. These shifts reflect the declining importance of physical sex differences due to (a) lower birthrates and much less reliance on lactation for feeding infants and young children, and (b) decreased reliance on strength and size as prerequisites for carrying out economically productive activities. These fundamental changes have set in motion far-reaching political, social, and psychological changes that have given women access to a greater range of social roles, including increased access to roles that yield authority and resources.

Despite the less extreme division of labor in contemporary industrialized societies, many sex differences remain. Women perform more domestic work than men and spend fewer hours in paid employment (US Bureau of Labor Statistics, 2008a). Although most women in the US are employed, they have lower wages than men, are concentrated in different occupations, and are rare at the highest levels of corporations and governments (Blau et al., 2006a; US Bureau of Labor Statistics, 2010). This division of activity yields less overall power, status, and resources for women than men (Valian, 1998), thereby retaining some degree of gender hierarchy or patriarchy. In our theory, these features of social structure – the division of labor and gender hierarchy – are the middle-level causes of sex-differentiated behavior (see [Figure 49.1](#)).

Gender Roles

Gender roles derive from the specific family and employment roles commonly held by women versus men in a society. Consistent with the correspondent inference principle (Gilbert and Malone, 1995), people infer the traits of men and women from observations of their behavior. Given a homemaker–provider division of labor, people disproportionately observe women and girls engaging in domestic behaviors such as childcare, cooking, and sewing, and men and boys engaging in activities that are marketable in the paid economy. Furthermore, perceivers tend to essentialize gender by viewing the different behaviors of the sexes as due to inherent differences in the natures of men and women. Thus, even though the division of labor is tailored to local conditions, it tends to be viewed by the members of a society as inevitable and natural.

The social behaviors that typify the home-maker-provider division of labor differ in their emphasis on communion versus agency (Eagly, 1987; Eagly and Steffen, 1984). Thus, women's accommodation to the domestic role fosters a pattern of interpersonally facilitative and friendly behaviors that can be termed communal. Women's communal activities encompass child-rearing, a responsibility that requires nurturant behaviors. The importance of close relationships to women's nurturing role favors the acquisition of superior relational skills and the ability to communicate nonverbally. In contrast, men's accommodation to the employment role, especially to male-dominated occupations, favors a pattern of relatively assertive behaviors that can be termed agentic (Eagly and Steffen, 1984).

The distribution of the sexes into occupations is another important source of observations of women and men. Given the moderately [p. 467 ↓] strong sex segregation of the labor force (Tomaskovic-Devey et al., 2006), perceivers infer the typical qualities of the sexes in part from observations of the type of paid work that they commonly undertake. Research has shown that occupational success is perceived to follow from agentic personal qualities to the extent that occupations are male-dominated and from communal personal qualities to the extent that they are female-dominated (Cejka and Eagly, 1999; Glick, 1991). Also, men have greater access to employment roles yielding higher levels of authority and income, and their adjustment to this aspect of their roles may foster relatively dominant behavior (Ridgeway and Bourg, 2004; Wood and Karten, 1986). Women's lesser access to such roles may favor more supportive behavior (e.g., Conway et al., 1996).

Gender roles are an important part of the culture and social structure of every society. Their power to influence behavior derives from their essential quality, appearing to reflect inherent attributes of women and men and from the related tendency to be relatively consensual and for people to be aware of this consensus (Wood and Eagly, 2010). Because gender roles are shared, people correctly believe that others are likely to react more approvingly to behavior that is consistent rather than inconsistent with these roles. Therefore, the most likely route to a smoothly functioning social interaction is to behave consistently with one's gender role or at least to avoid strongly deviating from it.

In summary, gender roles are emergent from the activities carried out by individuals of each sex in their typical occupational and family roles. To the extent that women more than men occupy roles that are facilitated by predominantly communal behaviors, domestic behaviors, or subordinate behaviors, corresponding attributes become stereotypic of women and part of the female gender role. To the extent that men more than women occupy roles that require predominantly agentic behaviors, resource acquisition behaviors, or dominant behaviors, the corresponding attributes become stereotypic of men and part of the male gender role. These gender roles, which are an important focus of socialization, begin to be acquired early in childhood and are elaborated throughout childhood and adolescence (e.g., Bussey and Bandura, 1999; Miller et al., 2006).

Gender Roles' Influence on Behavior

How do gender roles influence behavior? As Wood and Eagly (2010) argued, gender roles work through a trio of biosocial mechanisms to influence behavior in role-appropriate directions. These proximal causes of male and female behavior include biological processes involving hormonal changes and sociocultural factors of gender identity and others' stereotypic expectations. These three factors interact to yield both gender differences and similarities.

Influence of Hormonal Processes

Gender roles and specific social roles guide behavior in part through the activation of hormonal changes, especially in testosterone and oxytocin (Wood and Eagly, 2010). Hormones and related neural structures were shaped in part through ancient selection pressures associated with the basic perceptual, sensory, and motivational processes that humans share with other animals. In the standard interpretation, such inherited biological factors constrain sociocultural influences on men and women. More impressive is the recently emerging evidence that humans activate biological processes to support the sociocultural factors that guide masculine and feminine behaviors within cultures (Wood & Eagly, 2010). As these processes occur, subcortical

structures interact with more recently evolved, general-purpose, higher brain functions associated with the neocortex (Panksepp & Panksepp, 2000). Especially important to social interaction are processes located in the medial prefrontal cortex and the ventral anterior cingulate, which allow people to respond flexibly to others' expectations and self-regulate in response to their own identities [p. 468 ↓] (Heatherington, 2011). The evolution of the brain is thus a crucial component of our evolutionary theory, which stresses the importance of higher-level mechanisms for learning and innovation that are centered in the neocortex.

Biological processes include the activation of hormones that support culturally masculine or feminine behaviors. With culturally masculine roles, higher levels of testosterone are associated with dominance behaviors designed to gain or maintain status. In humans, such behaviors frequently entail competition, risk-taking, and aggression that may harm or injure others (Booth et al., 2006). In contrast, with culturally feminine roles, higher levels of oxytocin (as well as reduced cortisol and testosterone) are associated with behaviors that produce parental bonding, nurturance, and intimacy (Campbell, 2008).

Men and women selectively recruit hormones and other neurochemical processes for appropriate roles, in the context of their gender identities and others' expectations for role performance. Testosterone is especially relevant when, due to personal identities and social expectancies, people experience social interactions as dominance contests. Oxytocin is relevant when, due to personal identities and social expectancies, people define social interactions as involving bonding and affiliation with close others. Also important to bonding and affiliation are the neurochemical processes associated with rewards and learning of affiliation, which supplement or even supplant influences of oxytocin.

Influence of Gender Identities

Gender roles influence people's self-concepts and thereby become *gender identities* – individuals' sense of themselves as female or male. These identities arise because most people accept, or internalize, at least some aspects of cultural meanings associated with their sex (see Wood and Eagly, 2009, 2010).

People differ in the extent to which they incorporate gender roles into their self-concepts. Also, people differ in the aspects of gender roles that they adopt. For example, women who regard themselves as feminine could be invested in culturally feminine traits such as warmth and niceness or in feminine interests such as sewing or home decoration. People raised in culturally atypical environments may not internalize conventional gender-role norms. Consistent with research showing substantial relationships between sex-differentiated behaviors and self-reported agency and communion (Taylor and Hall, 1982), people who have self-concepts that differ from those that are typical of their sex are less likely to show gender-stereotypic behavior.

Gender identities motivate responding through self-regulatory processes. That is, people use their gender identity as a standard against which to regulate their behavior (Witt and Wood, 2010; Wood et al., 1997). People who have a masculine self-concept involving traits of dominance and assertiveness might regulate their behavior by, for example, seeking opportunities for leadership. Self-regulation proceeds in stages, beginning with *testing* the extent to which current behavior matches self-standards (e.g., Carver and Scheier, 2008). Closer matches produce positive emotions and increased self-esteem, whereas mismatches produce negative emotions and decreased esteem. When signaled by negative feelings, people *operate* on their behavior to bring it more in line with their desired standard. In this way, esteem and emotions constitute feedback about whether adjustments are necessary to meet standards.

Influence of Others' Expectations

A key assumption of our analysis is that both women and men typically are rewarded by other people for conforming to gender roles and penalized for deviating. Behavior consistent with gender role beliefs garner approval and continued interaction. In contrast, behavior inconsistent with gender roles is often negatively sanctioned and tends to disrupt social interaction. The sanctions for role-inconsistent behavior may be overt (e.g., losing a job) or subtle (e.g., being ignored, receiving disapproving looks).

[p. 469 ↓]

Research has produced abundant evidence of negative reactions to deviations from gender roles. For example, in a meta-analysis of 61 experiments on evaluations of male and female leaders, Eagly et al. (1992) showed that women who adopted a male-stereotypic assertive and directive leadership style were evaluated more negatively than men who adopted the exact same style, whereas women and men who adopted more democratic and participative styles were evaluated equivalently. In small-group interaction, women who behave in a dominant or extremely competent manner tend to lose likability and influence (Carli, 2001; Shackelford et al., 1996). Women in supervisory roles may be penalized for not attending to others' emotions or for expressing angry emotions (Byron, 2007; Brescoll and Uhlmann, 2008) as well as for performing at extremely well in stereotypically masculine roles (Heilman et al., 2004). In contrast, men may be penalized for behaving passively and unassertively (e.g., Anderson et al., 2001). People thus elicit conformity to gender-role norms by dispensing rewards such as liking and cooperation in return for conformity to these norms and dispensing social punishments such as rejection and neglect in return for nonconformity.

Because people often sanction behavior that is inconsistent with gender roles, these roles have a conservative impact by exacting costs from men and women who deviate from the norm. Given that men and women are aware of these costs, they are not likely to deviate from their gender role unless the behavior produced benefits that would outweigh the costs. Part of these perceived benefits for women, as members of a subordinate group in society, may be having a chance to gain access to rewards and opportunities usually reserved for men.

Behavior Influenced by Gender Roles and Specific Roles

Gender roles, as shared beliefs about men's and women's attributes, coexist with specific roles defined by factors such as family relationships (e.g., mother, son) and occupation (e.g., secretary, firefighter). In workplace settings, for example, a manager or lawyer occupies a role defined by occupation but is simultaneously a man or women and thus to some extent subjected to the constraints of his or her gender role. Similarly,

in a community organization, an individual who has the role of volunteer simultaneously is categorized as a woman or man and is thus perceived in terms of the expectations that are applied to people of that sex.

Because specific roles have direct implications for task performance in many natural settings, they can be more important than gender roles. This conclusion was foreshadowed by experimental demonstrations that stereotypic sex differences can be eliminated by providing information that specifically counters gender-based expectations. For example, Wood and Karten (1986) manipulated perceptions of agency in mixed-sex groups through false feedback that described participants as relatively agentic or not agentic. Controlling agency in this manner eliminated the usual sex differences in interaction style by which men, compared with women, showed more active task behavior and less positive social interactive behavior.

A field study by Moskowitz et al. (1994) used behavioral measures to examine the simultaneous influence of gender roles and organizational roles with a sample of adults who held a wide range of jobs in a variety of organizational settings. Implementing an experience-sampling method, this study found that agentic behavior was controlled by the relative status of the interaction partners, with participants behaving most agentically with a supervisee and least agentically with a boss. Yet, communal behaviors were influenced by the sex of participants, with women behaving more communally, especially in interactions with other women. Similarly, research on physicians demonstrated women's more communal behavior, even in the presence of a constraining occupational role. Female physicians, compared [p. 470 ↓] with male physicians, thus engaged in more partnership building with the patient, asked more questions, referenced more emotional and positive concerns, and offered more psychosocial information (e.g., concerning personal habits, impact on family; Roter et al., 2002).

It appears that employment roles provide relatively clear-cut rules about the performance of particular tasks. For example, regardless of whether a physician is male or female, he or she must obtain information about symptoms from a patient, provide a diagnosis, and design treatments to alleviate the patient's symptoms. Within the task rules that regulate physician–patient interactions, there is still room for some variation in behavioral styles. Physicians may behave in a warm, caring manner that focuses on

producing a positive relationship or in a less personally responsive style that focuses more exclusively on information exchange and problem solving. The female gender role may foster the caring, communal behavior that has been observed especially in female physicians as well as the participative, team-building style that has been observed especially in female managers (Eagly and Johnson, 1990; Eagly et al., 2003b). Thus, gender roles may have their primary influence on discretionary behaviors that are not required by the occupational role, which may often be behaviors in the communal repertoire. Gender roles are still important even though they become a secondary, background influence in settings in which specific roles are of primary importance.

Applications to Social Issues

Especially since the advent of the second-wave feminist movement, gender equality has been an important societal goal for many people. Yet debates continue about whether equality will come about, even though much social policy has been designed to promote equality. Affirmative action programs facilitating women's entry into nontraditional roles have produced mixed reactions (Heilman and Haynes, 2005). Women entering male-dominated roles contend with cultural incongruity between people's beliefs about what it takes to excel in those roles and stereotypes about the attributes of women (e.g., Eagly and Karau, 2002). As a result, even highly qualified women may be judged to lack the attributes necessary for success. Yet, these beliefs are not inevitable. As women assume nontraditional roles, people may develop new beliefs about women's attributes, given that these beliefs in part reflect role performance.

With respect to changes in men's and women's roles, an important question addressed by our theory is whether men and women differ in their behavior due to intrinsic sex differences built in through evolution or merely due to the social environment. As we have explained, our biosocial theory has a more complex view than these two opposing positions. Sex differences and similarities take a variety of forms, depending on men's and women's roles in society, which in turn reflect the more distal factors of male and female physical attributes, in combination with socioeconomic and cultural conditions. Roles in turn affect behavior through the immediate, proximal causes of hormonal regulation, self-regulation by gender identities, and social regulation by others' sanctions and rewards. It follows that female and male psychology is not fixed but emerges from

interactions across multiple biological and sociocultural factors. The varying forms of this interaction depend on the division of labor within a society and the ways in which boys and girls are socialized into sex-typical roles.

The dramatic changes that have occurred in women's roles in recent decades reflect the loosening of biosocial restraints through sharp reductions in birth rates and length of lactation combined with a shift toward an occupational structure that favors brains over brawn. These changes, combined with women's increased education, qualify them for occupations with more status and income. However, this shift has so far produced only partial equality between the sexes. [p. 471 ↓] Men continue to dominate leadership roles, especially at higher levels (e.g., Helfat et al., 2006). Women continue to take responsibility for the majority of childcare and housework (e.g., Bianchi et al., 2006), even when both spouses are employed full-time (US Bureau of Labor Statistics, 2008).

The continuing wage and authority gaps in the workforce can be traced in part to women continuing to fill caretaking roles, especially childcare. Childcare roles that take women out of the labor force or reduce their employment to part-time lessen their training and experience. This reduction of human capital, compounded by job discrimination against mothers (e.g., Correll et al., 2007), lessens women's opportunities to attain jobs that offer high wages and substantial workplace authority (Polacheck, 2006).

Social role theory offers a way to understand both stability and change in gender roles and associated behavioral sex differences. The recent social changes involved shifts in the roles of women as they moved into new educational and employment arenas. Thus, we expect to find convergence in those male and female attributes that reflect this masculinizing of women's experiences. A number of findings are consistent with this prediction. For example, women report increasing importance of job attributes such as freedom, challenge, leadership, prestige, and power, thus becoming more similar to men (Konrad et al., 2000). The career plans of female university students have shown corresponding changes (Pryor et al., 2006). Also, meta-analyses found decreases over time in the tendencies of men to engage in riskier behavior than women (Byrnes et al., 1999) and to emerge more than women as leaders in small groups (Eagly and Karau, 1991). Also, women's self-reported agency has increased over time to become more similar to men's agency, whereas the sex difference in communion has been relatively

invariant (Twenge, 1997, 2001; although see Lueptow et al., 2001). This increasing gender similarity in traditionally masculine domains is consistent with women's growing labor force participation and lessening concentration on childcare and other domestic activities.

The convergence of the sexes on masculine attributes is sufficiently marked to be apparent to everyday observers of men and women. Diekman and Eagly (2000) showed that people believe that women and men have converged in their personality, cognitive, and physical characteristics during the past 50 years and will continue to converge during the next 50 years. This perceived convergence mainly took the form of women increasing in the qualities typically associated with men. These studies also showed that perceivers function like implicit role theorists by assuming that, because the roles of women and men have become more similar, their attributes converge.

Despite this evidence of social change, men in industrialized nations do not appear to be undergoing transitions in their daily activities comparable to those of women. Men have only modestly increased their contributions to childcare and domestic work (Bianchi et al., 2006) and have yet to enter in large numbers into caring professions and other typically female types of paid employment (Queneau, 2006). Given the logic of social role theory, men's attributes will shift to include more communal qualities to the extent that they enter female-typical roles by performing more family caring activities and holding more communally demanding occupations.

Is it possible for men to adopt more communal roles and develop more caring, warm, emotionally-expressive personalities? A reasoned answer requires knowledge of the biosocial roots of the role structure and the limits it may impose on role flexibility. As we have explained, these roots lay mainly in the ways that male size and strength and female reproductive activities interact with socioeconomic complexity. Contemporary changes in female reproductive activity and the demands of occupations have eased women's inroads into male-dominated professional and managerial occupations and increased their agency. Barriers to men taking on childcare and communally demanding [p. 472 ↓] occupations include lower monetary compensation of these occupations (England, 2006), beliefs that men are less well endowed with communal skills (Cejka and Eagly, 1999), and stigma associated with nontraditional roles such as stay-at-home dads (Brescoll and Uhlmann, 2005).

Women undertake childcare more than men in part because of the continuing efficiency of this activity for women. The energetic demands of bearing children and the health benefits of some months of breastfeeding can orient mothers away from continued paid employment and toward infant care. This arrangement is fostered by female socialization and societal beliefs that promote sex-typical role performance. Hormonal processes also may encourage mothers' childcare, as the cascading hormones of pregnancy and lactation support women's tending (Campbell, 2008; Kuzawa et al., 2010; Taylor, 2002). Paternal behavior can also be supported by hormonal processes, as fathers show hormonal accommodation to parenthood similar to that of women (Berg & Wynne-Edwards, 2001; Wynne-Edwards, 2001). New fathers and mothers thus show suppressed testosterone and other hormonal changes, presumably to facilitate caretaking. Caretaking is additionally facilitated in both sexes by neurochemical mechanisms of reward learning that can undergird nurturing of infants and young children (Broad et al., 2006; Depue and Morrone-Strupinsky, 2005). Fathering activities also are fostered by changing norms and attitudes in the US, especially among younger adults, who have become more accepting of men's equal participation in childcare (e.g., Milkie et al., 2002).

In general, change toward gender equality is slowed by societal ideologies and status beliefs that legitimize social inequalities on the basis of sex and other attributes (Ridgeway, 2006; Sidanius and Pratto, 1999). Despite evidence that the pace of change in gender equality has slowed on many attitudinal and behavioral indicators since the second half of the twentieth century (Blau et al., 2006b), women's attitudes and ideologies are more progressive than men's (e.g., Eagly et al., 2003a), and their political commitments and actions continue to speed social change (e.g., Dodson, 2006; Seguino, 2007). For those committed to gender equality, a challenge for the future is to understand the roots of role asymmetries. Such understanding could facilitate social policy that opens paths for both men and women to occupy a wider range of social roles.

Alice H.Eagly WendyWood

References

- Abelson, R. A variance explanation paradox: When a little is a lot . Psychological Bulletin , 97, 129–133. (1985) <http://dx.doi.org/10.1037/0033-2909.97.1.129>
- Allport, G.W. (1954) The Nature of Prejudice . Cambridge, MA: Perseus Books.
- AndersonC., John, O.P., Keltner, D. and Krings, A.M. Who attains social status? Effects of personality and physical attractiveness in social groups . Journal of Personality and Social Psychology , 81, 116–132. (2001) <http://dx.doi.org/10.1037/0022-3514.81.1.116>
- Archer, J. Testosterone and human aggression: An evaluation of the challenge hypothesis . Neuroscience and Biobehavioral Reviews , 30, 319–345. (2006) <http://dx.doi.org/10.1016/j.neubiorev.2004.12.007>
- Bakan, D. (1966) The Duality of Human Existence: Isolation and Communion in Western Man . Boston: Beacon Press.
- Bem, S.L. The measurement of psychological androgyny . Journal of Consulting and Clinical Psychology , 42, 155–162. (1974) <http://dx.doi.org/10.1037/h0036215>
- Berg, S.J. and Wynne-Edwards, K.E. Changes in testosterone, cortisol, and estradiol levels in men becoming fathers . Mayo Clinic Proceedings , 76, 582–592. (2001)
- Berger, J., Rosenholtz, S.J. and Zelditch, M. Status organizing processes . Annual Review of Sociology , 6, 479–508. (1980) <http://dx.doi.org/10.1146/annurev.so.06.080180.002403>
- Bianchi, S.M., Robinson, J.P. and Milkie, M.A. (2006) Changing Rhythms of American Family Life . New York: Russell Sage.
- Biddle, B.J. (1979) Role Theory: Expectancies, Identities, and Behaviors . New York: Academic Press.

Blau, F.D., Brinton, M.C. and Grusky, D.B. (2006a) The declining significance of gender . In F.D. Blau, ed. , M.C. Brinton, ed. and D.B. Grusky (eds), *The Declining Significance of Gender?* pp. 215–244. New York: Russell Sage Foundation.

Blau, F.D., ed. , Brinton, M.C., ed. and Grusky, D.B. (eds) (2006b). *The Declining Significance of Gender?* New York: Russell Sage Foundation.

Booth, A., Granger, D.A., Mazur, A. and Kivlighan, K.T. Testosterone and social behavior . *Social Forces* , 85, 167–191. (2006) <http://dx.doi.org/10.1353/sof.2006.0116>

Brescoll, V.L. and Uhlmann, E.L. Attitudes toward traditional and non-traditional parents . *Psychology of Women Quarterly* , 35, 436–445. (2005) <http://dx.doi.org/10.1037/0003-066X.35.2.151>

Brescoll, V.L. and Uhlmann, E.L. Can an angry woman get ahead? Status conferral, gender, and expression of emotion in the workplace . *Psychological Science* , 19, 268–275. (2008) <http://dx.doi.org/10.1111/j.1467-9280.2008.02079.x>

Broad, K.D., Curley, J.P. and Keverne, E.B. Mother-infant bonding and the evolution of mammalian social relationships . *Philosophical Transactions of the Royal Society* , 361, 2199–2214. (2006) <http://dx.doi.org/10.1098/rstb.2006.1940>

Broverman, I.K., Vogel, S.R., Broverman, D.M., Clarkson, F.E. and Rosenkrantz, P.S. Sex-roles stereotypes: A current appraisal . *Journal of Social Issues* , 28, 59–78. (1972) <http://dx.doi.org/10.1111/j.1540-4560.1972.tb00018.x>

Buss, D.M. Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures . *Behavioral and Brain Sciences* , 12, 1–49. (1989) <http://dx.doi.org/10.1017/S0140525X00023992>

Buss, D.M. and Schmitt, D.P. Sexual strategies theory: An evolutionary perspective on human mating . *Psychological Review* , 100, 204–232. (1993) <http://dx.doi.org/10.1037/0033-295X.100.2.204>

Bussey, K. and Bandura, A. Social cognitive theory of gender development and differentiation . *Psychological Review* , 106, 676–713. (1999) <http://dx.doi.org/10.1037/0033-295X.106.4.676>

Byrnes, J.P., Miller, D.C. and Schafer, W.D. Gender differences in risk taking: A meta-analysis . *Psychological Bulletin* , 125, 367–383. (1999) <http://dx.doi.org/10.1037/0033-2909.125.3.367>

Byron, K. Male and female managers' ability to 'read' emotions: Relationships with supervisor's performance ratings and subordinates' satisfaction ratings . *Journal of Occupational and Organizational Psychology* , 80, 713–733. (2007) <http://dx.doi.org/10.1348/096317907X174349>

Campbell, A. Attachment, aggression and affiliation: The role of oxytocin in female social behavior . *Biological Psychology* , 77, 1–10. (2008) [http://dx.doi.org/10.1016/j.bipsycho.2007.09.001](http://dx.doi.org/10.1016/j.biopsycho.2007.09.001)

Carli, L.L. Gender and social influence . *Journal of Social Issues* , 57, 725–741. (2001) <http://dx.doi.org/10.1111/0022-4537.00238>

Carver, C.S. and Scheier, M.F. (1981) Attention and Self-regulation: A Control-theory Approach to Human Behavior . New York: Springer-Verlag. <http://dx.doi.org/10.1007/978-1-4612-5887-2>

Carver, C.S. and Scheier, M.F. (2008) Self-regulatory systems: Action and affect . In J.Y. Shah, ed. and W.L. Gardner (eds), *Handbook of Motivation Science* , pp. 308–324. New York: Guilford Press.

Cejka, M.A. and Eagly, A.H. Gender-stereotypic images of occupations correspond to the sex segregation of employment . *Personality and Social Psychology Bulletin* , 25, 413–423. (1999) <http://dx.doi.org/10.1177/0146167299025004002>

Conway, M., Pizzamiglio, M.T. and Mount, L. Status, communality, and agency: Implications for stereotypes of gender and other groups . *Journal of Personality and Social Psychology* , 71, 25–38. (1996) <http://dx.doi.org/10.1037/0022-3514.71.1.25>

Correll, S.J., Benard, S. and Paik, I. Getting a job: Is there a motherhood penalty? *American Journal of Sociology* , 112, 1297–1338. (2007) <http://dx.doi.org/10.1086/511799>

Cronbach, L. and Meehl, P. Construct validity in psychological tests . *Psychological Bulletin* , 52, 281–302. (1955) <http://dx.doi.org/10.1037/h0040957>

Deaux, K. (1976) *The Behavior of Women and Men* . Monterey, CA: Brooks/Cole.

Depue, R.A. and Morrone-Strupinsky, J.V. A neurobehavioral model of affiliative bonding: Implications for conceptualizing a human trait of affiliation . *Behavioral and Brain Sciences* , 28, 313–395. (2005)

Diekman, A.B. and Eagly, A.H. Stereotypes as dynamic constructs: Women and men of the past, present, and future . *Personality and Social Psychology Bulletin* , 26, 1171–1188. (2000) <http://dx.doi.org/10.1177/0146167200262001>

Dodson, D.L. (2006) *The Impact of Women in Congress* . New York: Oxford University Press. <http://dx.doi.org/10.1093/0198296746.001.0001>

Eagly, A.H. Sex differences in influenceability . *Psychological Bulletin* , 85, 86–116. (1978) <http://dx.doi.org/10.1037/0033-2909.85.1.86>

Eagly, A.H. (1987) *Sex Differences in Social Behavior: A Social-role Interpretation* . Hillsdale, NJ: Lawrence Erlbaum.

Eagly, A.H. and Carli, L.L. Sex of researchers and sex-typed communications as determinants of sex differences in influenceability: A meta-analysis of social influence studies . *Psychological Bulletin* , 90, 1–20. (1981) <http://dx.doi.org/10.1037/0033-2909.90.1.1>

Eagly, A.H. and Crowley, M. Gender and helping behavior: A meta-analytic review of the social psychological literature . *Psychological Bulletin* , 100, 283–308. (1986) <http://dx.doi.org/10.1037/0033-2909.100.3.283>

Eagly, A.H., Diekman, A.B., Schneider, M. and Kulesa, P. Experimental tests of an attitudinal theory of the gender gap in voting . *Personality and Social Psychology Bulletin* , 29, 1245–1258. (2003a) <http://dx.doi.org/10.1177/0146167203255765>

Eagly, A.H., Johannesen-Schmidt, M.C. and van Engen, M.L. Transformational, transactional, and laissez-faire leadership styles: A meta-analysis comparing women and men . *Psychological Bulletin* , 129, 569–591. (2003b) <http://dx.doi.org/10.1037/0033-2909.129.4.569>

Eagly, A.H. and Johnson, B.T. Gender and leadership style: A meta-analysis . *Psychological Bulletin* , 108, 233–256. (1990) <http://dx.doi.org/10.1037/0033-2909.108.2.233>

Eagly, A.H. and Karau, S.J. Gender and the emergence of leaders: A meta-analysis . *Journal of Personality and Social Psychology* , 60, 685–710. (1991) <http://dx.doi.org/10.1037/0022-3514.60.5.685>

Eagly, A.H. and Karau, S.J. Role congruity theory of prejudice toward female leaders . *Psychological Review* , 109, 573–598. (2002) <http://dx.doi.org/10.1037/0033-295X.109.3.573>

Eagly, A.H., Makhijani, M.G. and Klonsky, B.G. Gender and the evaluation of leaders: A meta-analysis . *Psychological Bulletin* , 111, 3–22. (1992) <http://dx.doi.org/10.1037/0033-2909.111.1.3>

Eagly, A.H. and Steffen, V.J. Gender stereotypes stem from the distribution of women and men into social roles . *Journal of Personality and Social Psychology* , 46, 735–754. (1984) <http://dx.doi.org/10.1037/0022-3514.46.4.735>

Eagly, A.H. and Steffen, V.J. Gender and aggressive behavior: A meta-analytic review of the social psychological literature . *Psychological Bulletin* , 100, 309–330. (1986) <http://dx.doi.org/10.1037/0033-2909.100.3.309>

Eagly, A.H. and Wood, W. Inferred sex differences in status as a determinant of gender stereotypes about social influence . *Journal of Personality and Social Psychology* , 43, 915–928. (1982) <http://dx.doi.org/10.1037/0022-3514.43.5.915>

Eagly, A.H. and Wood, W. Explaining sex differences in social behavior: A meta-analytic perspective . *Personality and Social Psychology Bulletin* , 17, 306–315. (1991) <http://dx.doi.org/10.1177/0146167291173011>

Eagly, A.H. and Wood, W. The origins of sex differences in human behavior: Evolved dispositions versus social roles . *American Psychologist* , 54, 408–423. (1999) <http://dx.doi.org/10.1037/0003-066X.54.6.408>

Eagly, A.H., Wood, W. and Fishbaugh, L. Sex differences in conformity: Surveillance by the group as a determinant of male nonconformity . *Journal of Personality and Social Psychology* , 40, 384–394. (1981) <http://dx.doi.org/10.1037/0022-3514.40.2.384>

England, P. (2006) Toward gender equality: Progress and bottlenecks . In F.D. Blau, ed. , M.C. Brinton, ed. , and D.B. Grusky (eds), *The Declining Significance of Gender?* pp. 245–264. New York: Russell Sage Foundation.

Friedan, B. (1963) *The Feminine Mystique* . Harmondsworth: Penguin

Gawronski, B. On difficult questions and evident answers: Dispositional inference from role-constrained behavior . *Personality and Social Psychology Bulletin* , 29, 1459–1475. (2003) <http://dx.doi.org/10.1177/0146167203256375>

Gilbert, D.T. and Malone, P.S. The correspondence bias . *Psychological Bulletin* , 117, 21–38. (1995). <http://dx.doi.org/10.1037/0033-2909.117.1.21>

Glass, A., McGraw, B. and Smith, M. (1981) *Integration of Research Studies: Meta-analysis of Research* . Beverly Hills, CA: Sage.

Glick, P. Trait-based and sex-based discrimination in occupational prestige, occupational salary, and hiring . *Sex Roles* , 25, 351–378. (1991) <http://dx.doi.org/10.1007/BF00289761>

Goffman, E. (1959) *The Presentation of Self in Everyday Life* . Garden City, NY: Doubleday Anchor.

Grossman, M. and Wood, W. Sex differences in intensity of emotional experience: A social role interpretation . *Journal of Personality and Social Psychology* , 65, 1010–1022. (1993) <http://dx.doi.org/10.1037/0022-3514.65.5.1010>

Hall, J.A. and Carter, J.D. Gender-stereotype accuracy as an individual difference . *Journal of Personality and Social Psychology* , 77, 350–359. (1999) <http://dx.doi.org/10.1037/0022-3514.77.2.350>

Hayden, B., Deal, M., Cannon, A. and Casey, J. Ecological determinants of women's status among hunter/gatherers . *Human Evolution* , 1, 449–473. (1986) <http://dx.doi.org/10.1007/BF02436620>

Heatherington, T.F. Neuroscience of self and self-regulation . *Annual Review of Psychology* , 62, 363–390. doi: .1146/annurev.psych.121208.131616. (2011) <http://dx.doi.org/10.1146/annurev.psych.121208.131616>

Heider, F. (1958) *The Psychology of Interpersonal Relations* . New York: Wiley. <http://dx.doi.org/10.1037/10628-000>

Heilman, M.E. and Haynes, M.C. (2005) Combating organizational discrimination: Some unintended consequences . In R.L. Dipboye, ed. and A. Colella (eds), *Discrimination at Work: The Psychological and Organizational Bases* , pp. 339–362. Mahwah, NJ: Erlbaum.

Heilman, M.E., Wallen, A.S., Fuchs, D. and Tamkins, M.M. Penalties for success: Reactions to women who succeed at male gender-typed tasks . *Journal of Applied Psychology* , 89, 416–427. (2004) <http://dx.doi.org/10.1037/0021-9010.89.3.416>

Helfat, C.E., Harris, D. and Wolfson, P.J. The pipeline to the top: Women and men in the top executive ranks of U.S. corporations . *Academy of Management Perspectives* , 20, 42–64. (2006) <http://dx.doi.org/10.5465/AMP.2006.23270306>

Hines, M. (2009) Gonadal hormones and sexual differentiation of human brain and behavior . In (eds), *Hormones, Brain, and Behaviour*, Vol. 4 , 2nd Edition, pp. 1869–1909. Oxford: Elsevier. <http://dx.doi.org/10.1016/B978-008088783-8.00059-0>

Hollingworth, L.S. Sex differences in mental traits . Psychological Bulletin , 13, 377–384. (1916) <http://dx.doi.org/10.1037/h0072261>

Horney, K. (1967) Feminine Psychology . New York: Norton.

Huber, J. (2007) On the Origins of Gender Inequality . Boulder, CO: Paradigm Publishers.

Konrad, A.M., Ritchie, J.E., Jr., Lieb, P. and Corrigall, E. Sex differences and similarities in job attribute preferences: A meta-analysis . Psychological Bulletin , 126, 593–641. (2000). <http://dx.doi.org/10.1037/0033-2909.126.4.593>

Kuzawa, C.W., Gettler, L.T., Huang, Y. and McDade, T.W. Mothers have lower testosterone than non-mothers: Evidence from the Philippines . Hormones and Behavior , 57, 441–447. (2010) <http://dx.doi.org/10.1016/j.yhbeh.2010.01.014>

Lueptow, L.B., Garovich-Szabo, L. and Lueptow, M.B. Social change and the persistence of sex-typing: 1974–1997 . Social Forces , 80, 1–36. (2001) <http://dx.doi.org/10.1353/sof.2001.0077>

Maccoby, E.E. (1966) Sex differences in intellectual functioning . In E.E. Maccoby (ed.), The Development of Sex Differences , pp. 25–55. Stanford, CA: Stanford University Press.

McKee, J.P. and Sherriffs, A.C. The differential evaluation of males and females . Journal of Personality , 25, 356–371. (1957) <http://dx.doi.org/10.1111/j.1467-6494.1957.tb01533.x>

Milkie, M.A., Bianchi, S.M., Mattingly, M.J. and Robinson, J.P. Gendered division of childrearing: Ideals realities, and the relationship to parental well-being . Sex-Roles , 47, 21–38. (2002) <http://dx.doi.org/10.1023/A:1020627602889>

Miller, C.F., Trautner, H.M. and Ruble, D.N. (2006) The role of gender stereotypes in children's preferences and behavior . In L. Balter, ed. and C.S. Tamis-LeMonda (eds), Child Psychology: A Handbook of Contemporary Issues , 2nd Edition, pp. 293–323. New York: Psychology Press.

Moskowitz, D.S., Suh, E.J. and Desaulniers, J. Situational influences on gender differences in agency and communion . *Journal of Personality and Social Psychology* , 66, 753–761. (1994) <http://dx.doi.org/10.1037/0022-3514.66.4.753>

Murdock, G.P. and Provost, C. Factors in the division of labor by sex: A cross-cultural analysis . *Ethnology* , 13, 203–225. (1973) <http://dx.doi.org/10.2307/3773347>

Olson, J.M., Roes, N.J. and Zanna, M.P. (1996) Expectancies . In E.T. Higgins, ed. and A.W. Kruglanski (eds), *Social Psychology: Handbook of Basic Principles* , pp. 211–238. New York: Guilford Press.

Parsons, T. and Bales, R.F. (1955) *Family, Socialization and Interaction Process* . Glencoe, IL: Free Press.

Polacheck, S. (2006) How the life-cycle human-capital model explains why the gender wage gap narrowed . In F.D. Blau, ed. , M.C. Brinton, ed. , and D.B. Grusky (eds), *The Declining Significance of Gender?* pp. 102–124. New York: Russell Sage Foundation.

Pryor, J.H., Hurtado, S., Saenz, V.B., Santos, J.L. and Korn, W.S. (2006) *The American freshman: Forty year trends* . Los Angeles: Higher Education Research Institute, UCLA. Available at: <http://www.gseis.ucla.edu/heri/norms06.php>, accessed January 27, 2009.

Queneau, H. Is the long-term reduction in occupational sex segregation still continuing in the United States? *Social Science Journal* , 43, 681–688. (2006) <http://dx.doi.org/10.1016/j.soscij.2006.08.019>

Ridgeway, C.L. Nonconformity, competence, and influence in groups: A test of two theories . *American Sociological Review* , 46, 333–347. (1981) <http://dx.doi.org/10.2307/2095064>

Ridgeway, C.L. Dominance, performance and status in groups: A theoretical analysis . *Advances in Group Processes* , 1, 59–93. (1984)

Ridgeway, C.L. (2006) Gender as an organizing force in social relations: Implications for the future of inequality , pp. 265–287. in F.D. Blau, ed. , M.C. Brinton, ed. , and D.B.

Grusky (eds), The Declining Significance of Gender? pp. 245–264. New York: Russell Sage Foundation.

Ridgeway, C.L. and Bourg, C. (2004) Gender as status: An expectation states theory approach . In A.H. Eagly, ed. , A.E. Beall, ed. and R.J. Sternberg (eds), The Psychology of Gender , 2nd Edition, pp. 217–241. New York: Guilford Press.

Rosenthal, R. How often are our numbers wrong? American Psychologist , 33, 1005–1008. (1978) <http://dx.doi.org/10.1037/0003-066X.33.11.1005>

Rosenthal, R. and Rubin, D.B. Interpersonal expectancy effects: The first 345 studies . Behavior and Brain Sciences , 1, 377–415. (1978) <http://dx.doi.org/10.1017/S0140525X00075506>

Ross, L. (1977) The intuitive psychologist and his shortcomings . In L. Berkowitz (ed.), Advances in Experiential Social Psychology , 10, 173–220. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960357-3>

Roter, D.L., Hall, J.A. and Aoki, Y. Physician gender effects in medical communication: A meta-analytic review . Journal of the American Medical Association , 288, 756–764. (2002) <http://dx.doi.org/10.1001/jama.288.6.756>

Salzman, P.C. Is inequality universal? Current Anthropology , 40, 31–61. (1999) <http://dx.doi.org/10.1086/515800>

Seguino, S. Plus ça change? Evidence on global trends in gender norms and stereotypes . Feminist Economics , 13, 1–28. (2007) <http://dx.doi.org/10.1080/13545700601184880>

Shackelford, S., Wood, W. and Worchele, S. Behavioral styles and the influence of women in mixed-sex groups . Social Psychology Quarterly , 59, 284–293. (1996) <http://dx.doi.org/10.2307/2787024>

Sidanis, J. and Pratto, F. (1999) Social Dominance: An Intergroup Theory of Social Hierarchy and Oppression . New York: Cambridge University Press.

Skrypnek, B.J. and Snyder, M. On the self-perpetuating nature of stereotypes about women and men . *Journal of Experimental Social Psychology* , 18, 277–291. (1982) <http://dx.doi.org/10.1016/0022-1031%2882%2990054-3>

Spence, J.T. and Helmreich, R.L. (1978) Masculinity and Femininity: Their Psychological Dimensions, Correlates, and Antecedents . Austin: University of Texas Press.

Strodtbeck, F.L. and Mann, R.D. Sex role differentiation in jury deliberations . *Sociometry* , 19, 3–11. (1956) <http://dx.doi.org/10.2307/2786099>

Swim, J.K. Perceived versus meta-analytic effect sizes: An assessment of the accuracy of gender stereotypes . *Journal of Personality and Social Psychology* , 66, 21–36. (1994) <http://dx.doi.org/10.1037/0022-3514.66.1.21>

Taylor, M.C. and Hall, J.A. Psychological androgyny: Theories, methods, and conclusions . *Psychological Bulletin* , 92, 347–366. (1982) <http://dx.doi.org/10.1037/0033-2909.92.2.347>

Taylor, S.E. (2002) *The Tending Instinct: How Nurturing is Essential to Who We Are and How We Live* . New York: Holt.

Taylor, S.E., Klein, L.C., Lewis, B.P., Gruenewald, T.L., Gurung, R.A.R. and Updegraff, J.A. Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight . *Psychological Review* , 107, 411–429. (2000) <http://dx.doi.org/10.1037/0033-295X.107.3.411>

Tomaskovic-Devey, D., Zimmer, C., Stainback, K., Robinson, C., Taylor, T. and McTague, T. Documenting desegregation: Segregation in American workplaces by race, ethnicity, and sex, 1966–2003 . *American Sociological Review* , 71, 565–588. (2006) <http://dx.doi.org/10.1177/000312240607100403>

Twenge, J.M. Changes in masculine and feminine traits over time: A meta-analysis . *Sex Roles* , 36, 305–325. (1997) <http://dx.doi.org/10.1007/BF02766650>

Twenge, J.M. Changes in women's assertiveness in response to status and roles: A cross-temporal meta-analysis, 1931–1993 . *Journal of Personality and Social Psychology* , 81, 133–145. (2001) <http://dx.doi.org/10.1037/0022-3514.81.1.133>

Uleman, J.S., Saribay, S.A. and Gonzalez, C.M. Spontaneous inferences, implicit impressions, and implicit theories . *Annual Review of Psychology* , 59, 329–360. (2008) <http://dx.doi.org/10.1146/annurev.psych.59.103006.093707>

US Bureau of Labor Statistics (2008) News — Married parents' use of time, 2003–06 . Available at: <http://www.bls.gov/news.release/pdf/atus2.pdf>, accessed December 10, 2008.

US Bureau of Labor Statistics (2010) Women in the labor force: A databook . Available at: <http://www.bls.gov/cps/wlf-databook-2010.pdf>, accessed January 11, 2010.

Valian, V. (1998) Why So Slow? The Advancement of Women . Cambridge, MA: MIT Press.

Witt, M.G. and Wood, W. Self-regulation of behavior in everyday life . *Sex Roles* , 62, 635–646. doi: 10.1007/s11199-010-9761-y. (2010) <http://dx.doi.org/10.1007/s11199-010-9761-y>

Wood, W. Meta-analytic review of sex differences in group performance . *Psychological Bulletin* , 102, 53–71. (1987) <http://dx.doi.org/10.1037/0033-2909.102.1.53>

Wood, W., Christensen, P.N., Hebl, M.R. and Rothgerber, H. Conformity to sex-typed norms, affect, and the self-concept . *Journal of Personality and Social Psychology* , 73, 523–535. (1997) <http://dx.doi.org/10.1037/0022-3514.73.3.523>

Wood, W. and Eagly, A.H. A cross-cultural analysis of the behavior of women and men: Implications for the origins of sex differences . *Psychological Bulletin* , 128, 699–727. (2002) <http://dx.doi.org/10.1037/0033-2909.128.5.699>

Wood, W. and Eagly, A.H. (2010) Gender . In S. Fiske, ed. , D. Gilbert, ed. , and G. Lindzey (eds), *Handbook of Social Psychology* , 5th ed., 1, pp. 629–667. New York: Oxford University Press.

Wood, W. and Eagly, A.H. (2009) Gender identity . In M. Leary, ed. and R. Hoyle (eds), Handbook of Individual Differences , pp. 109–125. New York: Guilford Press.

Wood, W. and Karten, S.J. Sex differences in interaction style as a product of perceived sex differences in competence . Journal of Personality and Social Psychology , 50, 341–347. (1986) <http://dx.doi.org/10.1037/0022-3514.50.2.341>

Wood, W., Rhodes, N. and Whelan, M. Sex differences in positive well-being: A consideration of emotional style and marital status . Psychological Bulletin , 106, 249–264. (1989) <http://dx.doi.org/10.1037/0033-2909.106.2.249>

Wynne-Edwards, K.E. Hormonal changes in expectant fathers . Hormones and Behavior , 40, 139–145. (2001) <http://dx.doi.org/10.1006/hbeh.2001.1699>

Zanna, M.P. and Pack, S.J. On the self-fulfilling nature of apparent sex differences in behavior . Journal of Experimental Social Psychology , 11, 583–591. (1975) <http://dx.doi.org/10.1016/0022-1031%2875%2990009-8>

<http://dx.doi.org/10.4135/9781446249222.n49>

Handbook of Theories of Social Psychology

Social Representation Theory

Contributors: Patrick Rateau & Pascal Moliner & Jean-Claude Abric & Pascal Moliner

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "Social Representation Theory"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n50>

Print pages: 477-498

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n50>

[p. 477 ↓]

Chapter 50: Social Representation Theory

Abstract

As heir to a strong French sociological tradition, the theory of social representations, elaborated by Serge Moscovici in the beginning of the 1960s, has become one of the major theories in social psychology. Mainly European initially, it rapidly brought together a large number of researchers and practitioners worldwide, mainly in the field of social psychology, but also in all other social sciences. These researchers have seen this theory as a flexible conceptual framework that enables us to understand and explain the way individuals and groups elaborate, transform, and communicate their social reality. They have also found in this theory's different developments a vast set of methods and tools, directly applicable to the analysis of a wide range of social issues. Lending itself equally well to qualitative approaches as to experimental applications, studies have multiplied along different lines. Those aiming at making connections between sociorepresentational processes and other processes classically studied in the field of social cognition seem to be the most promising in terms of the theory's future development. This chapter addresses a longstanding tradition of research, covering a period of nearly 100 years of research, from 1893 to 2010.

Introduction

A Common Sense Theory

In many ways, social psychology is the study of social reality. That is to say that it deals with the explanations to which we automatically have recourse in order to explain and understand the world around us. Indeed, each one of us desires to make sense of

events, behaviors, ideas, and exchanges with others and seeks to find around them a certain coherence and stability. Each one of us seeks to explain and understand their environment in order to make it predictable and more controllable. Yet, this environment is made up of innumerable situations and events, and a multiplicity of individuals and groups. Similarly, we are being constantly required, during our everyday interactions, to make decisions, to give our opinion on this or that subject or to explain this or that behavior. In short, we are constantly plunged into an environment where we are bombarded with information and required to deal with it. In order to understand, [p. 478 ↓] master, and make sense of this environment we have to simplify it, to make it more predictable and familiar. In other words we have to reconstruct it in our own fashion.

But one cannot help but notice that this process of reconstruction is a constantly repeated process. From our youngest age, school, the family, institutions, and the media, instill in us certain ways of seeing the world and offer us a particular vision of the things around us, presenting us largely with a ready-made construction of the world in which we grow up, the values with which it is invested, the categories which govern it and the principles themselves by which we understand it. Our perception of the environment is next shaped by the groups, the associations, and the clubs that we become part of. It is very largely in our exchanges and our communications with others that our reality of the world around is formed. In the course of our contacts and our multiple involvements with different social groups we ourselves acquire and transmit knowledge, beliefs, and values that allow us to share a common conception of things and of others. In this sense, this reconstruction of reality, this representation of reality, is above all social; that is to say elaborated according to the social characteristics of the individual and shared by a group of other individuals having the same characteristics.

This last point is important. Not all social groups share the same values, the same standards, the same ideologies, or the same concrete experiences. Yet all construct representations that are closely based on these. It follows that social representations bear on the one hand the mark of the social membership of the individuals who adhere to them and give them their identity, and on the other allow these same individuals to distinguish "others," those who do not share the same representations and who appear to them at best as different, at worst as enemies.

To sum up, social representations can be defined as “systems of opinions, knowledge, and beliefs” particular to a culture, a social category, or a group with regard to objects in the social environment. At this introductory stage, it seems unnecessary to go any further. We will simply note at this point that with regard to social representations the distinction between the notions of “opinions,” “knowledge,” and “beliefs” is unnecessary. Of course opinions are mostly concerned with the field of position taking, knowledge with the field of learning, and experience and beliefs with that of conviction. But our everyday experience shows us that for individuals, there is frequently confusion between these three areas, especially when talking about a socially invested object. To this effect, we observe beliefs that have the status of established truths, or opinions that look peculiarly similar to beliefs, with the result that the lines between what “I think,” “I know,” and “I believe” often become blurred. As a consequence, the contents of a representation may be indifferently classed as opinions, information, or beliefs, and we may choose that a social representation comes across concretely as a set of “cognitive elements” relative to a social object.

The first characteristic of this set is that of *organization*. This is well and truly a structure, and not just a collection of cognitive elements. This means that the elements that constitute a social representation interact with each other. More exactly, this means that people cooperate in establishing relationships between these diverse elements. Particular opinions are considered equivalent to others, particular beliefs are deemed incompatible with particular information, and so on.

The second specificity of a representation is that of being *shared* by the members of a particular social group. However, the consensuses observed on the elements of a given representation depend at the same time on the homogeneity of the group and on its members' position towards the object, so that the consensual nature of a representation is generally partial, and localized to certain elements of the latter.

The third characteristic of this set resides in its method of construction; it is *collectively produced* through a more global process of communication. Exchange between individuals and exposure to mass communication [p. 479 ↓] allow the members of a group to share the elements that will constitute a social representation. This sharing process favors the emergence of a consensus at the same time as conferring social validity on diverse opinions, information, and beliefs.

Finally, the fourth specific role of a social representation concerns its purpose – it is *socially useful*. Firstly, of course, in order to understand the object to which the social representation refers. Representations are above all systems allowing the understanding and interpreting of the social environment. But they also intervene in interactions between groups, particularly when these interactions are engaged in around a social object. Every society, as shown by Adam Smith (1776) and Emile Durkheim (1893), revolves around the division of labor. This division is not only a condition of social cohesion, but also a permanent source of dependency and power relationships within the community. Indeed, it leads to the differentiation of groups, roles, status, professions, castes, and so on. Thus, everyone is interdependent whilst being clear about their separate identity. Complementarity and differentiation are two interdependent operations that are fully active within representations. Furthermore, social representations provide criteria for evaluating the social environment that enable determination, justification, or legitimization of certain behaviors. Taken together, that is how Serge Moscovici (1961) defines the notion of social representation in his first work devoted to the image and the dissemination of the psychoanalytic theory in France in the middle of the twentieth century. Whilst studying the way in which a scientific theory is transformed into a common sense theory, Moscovici traced the first outlines of what would henceforth be called the theory of social representations (SRT), whose success has not wavered since.

The Liveliness of the SRT

Some 50 years after its introduction to the field of social psychology, the importance of the SRT is well known; without doubt it is a major theoretical and empirical movement. The reasons for this success are diverse.

Let us start with its interdisciplinary nature. Located in the social and psychological interface, the social representation concept is of interest to all the social sciences. It has found a place in the fields of sociology, anthropology, history, geography, and economy and studies are carried out on its links with ideologies, symbolic systems, and attitudes. But it can also be found in the fields of cognition and linguistics. This multiplicity of relations with other disciplines confers on the SRT a transversal status that mobilizes

and connects different fields of research. This interdisciplinary nature constitutes without a doubt one of the most fertile and dynamic contributions made by this field of study.

The second reason is the flexibility of its conceptual framework which has enabled this theory to adapt to various research areas (communication, social practice, intergroup relations, etc.), and to initiate many theoretical and methodological developments. But to these reasons can be added another, more fundamental point from our perspective. As a “socially built and shared knowledge theory” (Jodelet, 1989), the SRT is a theory of social bonding. It gives us an insight on what permanently connects us to the world and to others. It teaches us about how this bond is built. In this sense, one can see here a global theory of the social individual and a possible way for integrating the diverse paradigms and fields of social psychology. The success of the SRT can be measured in terms of its scientific verve. Indeed, ever since the founding work by Serge Moscovici innumerable works have regularly presented new research developments in the field of social representations. In France, this phenomenon has been particularly marked since the 1980s when publications devoted to this theme appeared approximately every three years. It was also in the 1980s that the theory began its rapid expansion abroad, with the publication and translation into English of many books on the subject (Breakwell and [p. 480 ↓] Canter, 1979; Deaux and Philogène, 2001a, 2001b; Duveen, 2001; Farr and Moscovici, 1984; Moscovici, 1981, 1982, 1984, 1988, 2001a, 2001b; Mugny and Carugati, 1989).

According to the census conducted by Vergès (1996), the SRT, with more than 2000 articles, laid claim to be one of the most famous psychosocial theories, at the same level as cognitive dissonance, which, in its 27 years of existence had had more than 1,000 references (Cooper and Croyle, 1984). In addition, regular international symposiums are dedicated to it (Ravello, 1992; Rio de Janeiro, 1994; Aix-en-Provence, 1996; Mexico, 1998; Montréal, 2000; Stirling, 2002; Guadalajara, 2004; Rome, 2006; Bali, 2008; Tunis, 2010), as are many journals and special editions of journals. Finally, we should mention the creation of an Internet network (Social Representations and Communication Thematic Network) bringing together researchers worldwide (South America, the US, Japan, India, Russia, etc.) and of a European PhD on Communication and Social Representations in 1993. If one can say that a good theory is one that is “talked about,” then the sheer quantity of communication around the theory of social representations confers on it the status of a major theory.

Ultimately, the scientific assessment of the SRT may appear to be somewhat flattering. However, it has not always been like this. By examining the historical development of the SRT, we will attempt to show how it progressively found its place in the field of social psychology, the different orientations running through it at present, what connections it has with other major psychosocial paradigms, and finally, in what way it constitutes today an essential theory for analyzing and understanding social problems.

A Brief History of the Theory and its Developments

After having been the most memorable phenomenon in French social sciences at the beginning of the twentieth century, the notion of collective representations, introduced by the French sociologist Emile Durkheim in 1898, fell into disuse for more than 50 years. It was towards the beginning of the 1960s that Moscovici renewed studies of the concept and aroused the interest of a small group of social psychologists, thus bringing the theory back to life. They saw in it the possibility for tackling their discipline's issues from a new and original angle (Abric, 1976; Codol, 1970; Flament, 1971). The study of knowledge dissemination, of the relationship between thought and communication, and of the genesis of common sense, formed the elements of a new program that has been familiar ever since. But, in between the concept of collective representations and the contemporary researches on social representations, the concept has undergone many a metamorphosis, giving it different forms and colors. It is this history that we shall attempt to retrace here.

From Collective Representations to Social Representations

All attempts to reconstitute the past of this concept necessarily begin with sociology. Simmel (1908) was without doubt the first to recognize the connection between the separation of the individual who distances himself from others and the necessity to symbolize these others. He argued that the manner in which we symbolize others

shapes reciprocal action and the social circles that they form together. From a different point of view, Weber (1921) saw representations as a reference framework and a channel for action by the individual. He attempts to describe a common knowledge capable of anticipating and prescribing the behavior of individuals.

But the true inventor of this concept is Durkheim (1893, 1895, 1898) insofar as he defines its contours and recognizes its ability to explain various societal phenomena. He defines it as a double separation. First, collective representations are to be distinguished [p. 481 ↓] from individual representations. The latter, unique to each individual, are extremely variable, fleeting, short-lived and constitute a steady stream, whereas collective representations are impersonal and untouched by time. Second, individual representations are rooted in the individual consciousness, whereas collective representations are mutually held throughout society. Such representations are thus homogeneous and shared by all members of society. Their function is to preserve what binds them, to prepare and act in a uniform manner. This is why they are collective, why they are handed down over the years from generation to generation, and why they act for individuals as strong cognitive constraints. For Durkheim, the aim is clear: collective thinking has to be studied in itself and for itself. The forms and content of representations have to become a separate domain in order to be able to claim and prove social autonomy. For him, this is social psychology's task, even though it's still in its formative stages and its purpose still seems unclear.

However, during the very beginning of the twentieth century, it was above all sociology, anthropology, and ethnology (Lévi-Strauss, 1962; Lévy-Bruhl, 1922; Linton, 1945; Mauss, 1903) which would use the notion of representations, in a perfectly descriptive manner, to study different collective representations in cultural or ethnic communities. It was not until the 1960s that, following in Durkheim's footsteps, and based on child (Piaget, 1932) and clinical psychology (Freud, 1908, 1922), Serge Moscovici (1961) attempted to elaborate a social psychology of representations. Considering that Durkheim's conceptions left relatively little place for the question of interactions between the individual and the collective, he proposed to replace the term "collective representation" with a more restricted "social representation." In the words of the author, it was to

transfer to modern society a notion that seemed to be reserved to more traditional societies [in response to the] necessity of making representations into a bridge between individual and social spheres, by associating them with the perspective of a changing society. (Moscovici, 1989: 82)

This evolution is marked by two fundamental changes in relation to Durkheimian conceptions.

First, Moscovici considers that representations are not the product of society as a whole, but the products of the social groups who build this society. Second, he focuses on communication processes, considered as explaining the emergence and transmission of social representations. The first point allows the conception of a social mentality which is overdetermined by societal structures and also by the insertion of individuals in these structures, in such a way that different social representations of the same object are seen to exist within a given society. The second change to the representation theory, introduced by Moscovici, permits the conception that through communication – and the influence, normalization and conformity processes that go with it – individual beliefs can be the object of a consensus at the same time as collective beliefs can impose themselves on the individual.

However, the social representation concept would undergo another period of latency before mobilizing the broad stream of research mentioned earlier in this chapter. The theory's true deployment couldn't happen until many epistemological obstacles had been removed, the largest of them all being the behaviorist model, which denied any validity to the consideration of mental processes and their specificity. The decline of behaviorism and the emergence of the "new look" in the 1970s, followed by cognitivism in the 1980s, led to the progressive expansion of the "stimulus-response" (S-R) paradigm. This development meant that internal psychological states, conceived as an active cognitive construction of the environment and dependent on individual and social factors, were recognized as having a creative role in the behavior elaboration process. This is perfectly expressed by Moscovici, when he says that representations determine at the same time stimuli and responses, in other [p. 482 ↓] words "that there is no line between the external and internal universes of individuals or a group" (1969: 9).

This overturning of perspectives marked, from the 1980s, the development and improvement of work on social representations. It is also considered, in a diagrammatic sense, that these works were developed along three main lines, each one attempting to develop different facets of the concept. One which examines the regulatory role of social representations on real social interactions, another which studies the impact of social relationships on the elaboration of social representations, and a third which analyses representational dynamics and their structural characteristics, more specifically linked to social conduct. These three lines of development revolve not so much around different points of view as different ways of approaching social representations. This diversity of orientation most probably comes from the fact that Moscovici himself proposed diverse definitions of social representations, all of which are complementary.

There are multiple reasons for this flexibility. First of all, research is not limited by being enclosed within a rigid and narrow theoretical framework. Second, it allows the study of social representations to be situated within the framework of a paradigm, a line of thought and a knowledge structuring tool, rather than within an established and narrow-minded theoretical framework. Finally, the reality of social representations is such that their definition can vary according to the researcher's perspective. We can therefore study them in their emergence and in their role as regulator of social interaction and communications, from the angle of their internal structure or even from that of their links with social relations. We are now going to briefly introduce these three perspectives.

Orientations of the SRT

The Sociogenetic Model

When he formed his theory, Moscovici (1961) wanted above all to propose a description of the genesis and the development of social representations. According to him, the emergence of a social representation always coincides with the emergence of an unprecedented situation, unknown phenomenon, or unusual event. This new nature of the object implies that information about it is limited, incomplete, or widely spread

throughout the different social groups involved with the emergence of this object (what Moscovici called the *dispersion of information*). This new object arouses worry and vigilance or disrupts the normal course of things. It thus motivates intense cognitive activity to understand it, control it, or even defend oneself from it (*inference pressure* phenomenon), and causes a multiplicity of debates and of interpersonal and media communication. As a result of this information, beliefs, hypotheses, or speculations are shared, leading to the emergence of majority positions in different social groups. This emergence is facilitated by the fact that individuals deal with information on the object or the situation selectively, focusing on particular aspects according to their expectations and the orientations of the group (*focalization* phenomenon).

The gradual emergence of a representation occurs spontaneously and is based on three kinds of phenomena: the dispersion of information, focalization and the pressure to make inferences. But these phenomena themselves are developed on the basis of two major processes defined by Moscovici: *objectification* and *anchoring*.

Objectification refers to the way in which a new object, through communication about it, will be rapidly simplified, imaged, and diagrammed. Through the phenomenon of selective construction, different characteristics of the object are taken out of context and sorted according to cultural criteria (all groups do not have an equal access to object relative information), to normative criteria (only what agrees with the group's system of values is retained). The different aspects of the object are thus separated from the field to which they belong to be appropriated by groups who, by projecting them into their [p. 483 ↓] own reality, can control them more easily. These selected elements form together what Moscovici calls a *figurative core*, that is to say a coherent visualization that reproduces the object in a concrete and selective manner. By penetrating the social body through communication, by collective generalization, this simplification of the object replaces the object's reality, and is "naturalized." A representation is then created and takes on an "obvious" status. As such, it is an "independent theory" of the object which will serve as a basis for judgments and behavior oriented towards it.

To this effect, Moscovici, while studying the emergence of the representation of psychoanalysis in French society, observed the apparition of a figurative core composed of four parts: the conscious, the unconscious, repression, and complexes. These elements are fully extracted from their original theoretical context. They are also

naturalized in the sense that individuals don't consider them as abstract notions but as concrete and observable elements of the psychic apparatus. From there comes the possibility to communicate about psychoanalysis beyond its conceptual framework, to recognize categories of disorders and symptoms (the superiority complex, modesty, the slip, unconscious repression, subconscious acts, etc.) and different categories of people (the complicated, the repressed, the neurotic, etc.).

Anchoring completes the objectification process. It corresponds to the way an object finds its place in a pre-existing individual and group thought system. Depending on an elementary mode of knowledge production based on an analogy principle, the new object is assimilated into forms that are already known and into familiar categories, and so on. At the same time, it will become identified with a network of already present meanings. The hierarchy of values belonging to different groups constitutes a meaning network in which the object will be located and evaluated. The object will thus be interpreted in different ways depending on social groups. Furthermore, this interpretation extends to anything that remotely concerns this object. Thus, all social groups attach the object to their own meaning networks, guarantors of their identity. In this way a vast set of collective meaning is created around the object. In this way also, the object becomes a mediator and a criteria for relationships between groups. However, and this is an essential point to anchoring, integrating the new object into a pre-existent system of norms and values cannot happen smoothly. An innovative mix results from this contact with the new and the old, due both to the integration of the hitherto unknown object, and to the persistence of the old, the new object reactivating habitual frameworks of thought in order to incorporate it. From this it follows that a social representation always appears as innovative and enduring, changing and unchanging.

On this general theoretical basis of the process of generating social representations has developed a large research field, initiated notably by the work of Denise Jodelet (1989). This stream of research focuses on the descriptive study of social representations as meaning systems that express the relationships that individuals and groups have with their environment. Considering that representations are born essentially through interaction and contact with public discourses, this line of research concentrates firstly on language and speech from two complementary viewpoints. Social representations are approached as being at once fixed *in* language and as functioning themselves as

language through their symbolic value and the framework they supply for coding and categorizing individuals' environment.

So-called monographic and qualitative approaches to discourse and behavior data collection and analysis (ethnographic techniques, sociological investigations, historical analysis, in-depth interviewing, focus groups, discourse analysis, documentary analysis, verbal association techniques, etc.) constitute the main methodological framework for works carried out in this area (see, for example, Kronberger and Wagner, 2000; Markova, 1997, 2003; Wagner, 1994; Wagner et al., 1999).

[p. 484 ↓]

The Structural Model

Based at the same time on Moscovici's objectification process and on Asch's work on social perception (1946), Jean-Claude Abric and Claude Flament proposed an approach known as the "central core theory" (see Abric, 1993, 2001). This approach has massively contributed to clarifying sociocognitive logics underlying the general organization of social representations.

We are reminded that, at the time of his famous observations, Asch showed that amongst the seven character traits suggested to subjects as criteria for evaluating the image of a partner, one of them (warm/cold) played a principal and central role in the process studied, inasmuch as it played a far greater role in determining the perception of the other person than the other traits proposed.

Inspired by these results, Abric proposed transcending the purely genetic framework of the figurative core idea by recognizing its paramount role in all established representations. The basis of the central core theory is to consider that, in the overall picture of cognitive elements which make up a representation, certain elements play a different role to others. These elements, called central elements, form a structure named by Abric the "central core." This internal structure of representations provides two essential functions: (a) a meaning generative function – it is through the central core that other elements in the representational field acquire meaning and specific value for

individuals; and (b) an organizational function – it is around the central core that other representational elements are arranged. And it is this same core that determines the relationships that these elements maintain with each other.

Thus, as a cognitive structure providing meaning generative and organizational functions, the core structures in its turn elements that refer to the object of representation. These elements, dependent on the core, are called “peripheral elements.”

As proposed by Flament (1989), in reference to the scripts theory (Schank and Abelson, 1977), these peripheral elements allow representations to function as a “decryption” grid of social situations experienced by individuals. If the central core can be understood as the abstract part of the representation, the peripheral system should be understood as its concrete and operational one.

In the end, according to Abric, social representations act as entities, but with two different and complementary components:

- 1. The central system structures cognitive elements relative to an object and is the fruit of particular historical, symbolic, and social determinisms to which different social groups are subject. It is characterized by two fundamental properties. First, by a great stability, thus assuring the permanence and durability of the representation. In other words, the central system resists any scrutiny, in one way or another, of the representation's general basis. It is, moreover, where consensus on the representation is found, and thus constitutes its collectively shared common basis. It enables each group member to “see things” in approximately the same way, and through it, the group's homogeneity concerning the representation's object is defined. Thanks to the central system, group members can recognize each other, but also differentiate themselves from neighboring groups, and thus, to a great extent, it contributes to social identity.
- 2. The peripheral system, in tune with every day contingencies, enables a representation to be adapted to various social contexts. Flament assigns to it three essential functions:

- (a). It prescribes behavior and position taking allowing individuals to know what is and is not normal to say or do in a given situation, in view of its purpose.
- (b). It permits personalization of the representation and of the behaviors which are linked to it. Depending on the context, the same representation can lead to different interpersonal opinions within a group. These differences remain compatible with the central system, but correspond to an internal variability of the peripheral system.
- (c). It protects the central core when necessary and acts as a representation's "bumper." In this sense, the transformation of a social representation occurs in most cases through the prior modification of peripheral elements.

[p. 485 ↓]

From an epistemological point of view, the structural approach marks a major turning point for the theory of social representations. On the one hand, because it provides researchers with a conceptual framework for studying stabilized representations rather than representations in their formative stage. Seen from this perspective, social representations are no longer simple "spheres of opinions," but become structured spheres. In this sense, the study of their structure takes over from that of their content. On the other hand, the structural approach offers a framework for analysis which allows us to identify the interaction between the functioning of the individual and the contexts in which the individual evolves. Finally, because the structural approach offers formalized concepts, it allows the formulation of hypotheses around the sociocognitive adaptation of social actors faced with the evolutions of their environment. And these hypotheses are at the origin of the experimental method in the study of social representations.

The Sociodynamic Model

Based on the anchoring process defined by Moscovici, Willem Doise (see Clémence, 2001 for an overview) proposed a theoretical model which aimed to reconcile the structural complexity of social representations and their insertion in plural social and ideological contexts.

According to Doise, representations can only be envisaged in the social dynamic which, through communication, places social players in interactive situations. This social dynamic, when elaborated around important issues, arouses specific position taking, in relation to the social integration of individuals. That is to say that positions expressed on a given question depend fundamentally upon peoples' social memberships, which refers back to Moscovici's anchoring process. But Doise adds that these positions depend also on the situations within which they are produced. This double source of variation can generate an apparent multiplicity of position taking even though they arise from common organizational principles. Indeed, for Doise, all social interactions have symbolic characteristics. They enable people and groups to define themselves in relation to others. They therefore contribute to defining everyone's identity. This is why they have to be organized according to common rules among specific group members. By supplying shared "reference points" serving as a basis for the position taking of individuals and groups, representations constitute common rules. They thus organize the symbolic processes which underlie social interaction.

In other words, this model assigns a double role to representations. They are defined, firstly, as principles that generate position taking. But they are also principles for organizing individual differences. On the one hand, they supply individuals with common reference points. On the other hand, these reference points become issues that individual differences revolve around. If representations allow the object of the debate to be defined, they also organize this debate by suggesting the questions to be asked.

In this conception, there isn't necessarily a consensus regarding opinions expressed by individuals. It is not the points of view which are shared, rather it is the questions which attract conflicting points of view. To sum up, position taking can diverge even when referring to common principles. Let us note finally that the theory of organizational principles gives great importance to intergroup relationships, by trying to show how different social memberships can determine the importance given to different principles. From this perspective, it's to do with studying the anchoring of representations in collective realities.

The sociodynamic approach introduces a new way of thinking of the question of consensus in the SRT. For Moscovici, this consensus resulted from the sharing of certain beliefs within a given group. And this sharing was itself the result of the

communication process. Doise considered consensuses more as anchoring points for a social representation. And the convergences or divergences [p. 486 ↓] between these anchoring points find their origin in the structuring of existing social relations between groups. Seen from this perspective, the study of social representations needs to make use of multiple approaches that will highlight the links between cognitive elements and also between individuals or groups and cognitive elements (see Doise et al., 1992). So it is a question of establishing principles of homology between the social positions of individuals and their position taking in order to reveal the organizing principles of the representations studied (see Clémence, 2001; Lorenzi-Cioldi and Clémence, 2001; 2010; Spini, 2002).

The Expansion of the Theory

These three theoretical orientations developed by French and Swiss researchers constituted, and still constitute, the bases on which would develop, notably from the 1980s, a multitude of studies, first from outside of Europe, mainly in Latin America.

Very soon, and mainly under the influence of Robert Farr and Miles Hewstone, the SRT gained a foothold in the UK from which emerged, for example, the work of Gerard Duveen centering on the connection between the individual and the group within the framework of microgenetic socialization processes; that of Sandra Jovchelovitch who proposes the view of social representations as a space between the individual and the society linking objects, the subject, and activities; that of Caroline Howarth centering on the links between the SRT and social identity; or yet again that of Ivana Markova who is developing links between dialogicity and social representations. In Austria, the work of Wolfgang Wagner in particular has demonstrated the role between social interactions and discursive exchanges in the processes of construction of social representations. In Italy, under the impetus of Augusto Palmonari, then of Felice Carrugati, the work of Anna Maria de Rosa led to the establishment and dissemination of the SRT throughout Europe. On the other side of the Atlantic, it was mainly in the countries of Latin America and South America (particularly Mexico, Brazil, Argentina and Venezuela) that the SRT found, from the 1990s onwards, very fertile ground for expansion. The impact of social, historical, and cultural contexts on the formulation of Latin American scientific issues had a lot to do with this success. Researchers in social psychology have found in it

a creative, reflexive and critical way of thinking, suitable for dealing with change and political, economic, and social crises. They participate actively today in the theoretical developments of the SRT by linking it particularly with other psychosocial issues such as, for example, social memory or the processes of social change. We should also mention studies carried out in Portugal, Spain, and Rumania and more recently in Australia, Asia, and Africa, but one chapter does not give us adequate space to do so.

We will point out, on the other hand, that in this international picture, the US is one of the most notable absentees. Despite the remarkable work of Gina Philogene and Serge Moscovici to attempt to integrate the SRT into North American social psychology studies, one cannot but notice that it has not found true ground for development. The reasons for this are many and once again there is not enough space here to draw up a coherent and detailed list. The relative laxity of the initial theoretical arguments and the publication almost exclusively in French of the first developments in the SRT are undoubtedly among the main reasons. But there are also more profound and metatheoretical reasons which have long made SRT and social cognition strangers to each other.

Amongst these reasons, that which appears to us to have the most weight concerns the difference in the types of analysis assigned to research carried out in the two fields. Traditionally, social cognition is mainly interested in the intraindividual processes which underlie social interaction, whereas SRT is historically concerned with interindividual phenomena (Kruglanski, 2001), which affect the consciousness of the individual. [p. 487 ↓] The bridging of the gaps between these two fields of study constitutes without doubt one of the most fascinating scientific issues for the years to come in the field of social psychology. It is also in this direction that a part of our own work lies.

Personal Narrative of the Theory's Development

Our personal involvement in social representation research dates back to the mid-1980s. At the time, the theory was beginning to expand rapidly in France and in Europe, but was still the object of much criticism. The theory was reproached mainly for being

too flexible in terms of concepts and lacking in terms of methodology. Basically, what new aspects did social representations bring to the notions of opinion and attitude, already solidly anchored in social psychology? To answer this criticism, a team of researchers from the University of Aix en Provence proposed two arguments. For Jean Claude Abric and Claude Flament, who were leading this team, representations had to be conceived as cognitive structures. They were not just spheres of opinions, as advanced by Moscovici, but well and truly structured groups within which some elements had a specific role to play. Moreover, even if this idea wasn't yet clearly formed, Abric and Flament thought that, contrary to attitudes, essentially linked to the *evaluation* of social objects, representations concerned above anything else the *meaning* of those same objects. Basically, the idea was that it is the representation which defines the object of the attitude.

Based on these arguments, it was necessary to propose a theory that would account for both the structure and the dynamics of a social representation. This theory already existed, it had been proposed by Jean-Claude Abric in 1976. It still had to be confirmed and demonstrated that it allowed the stability and the dynamic of representations to be better described. It was in this context that two of us joined the Aix en Provence team as doctoral students. In 1988, two theses were defended. The first showed that within social representations, certain beliefs effectively play a specific role (Moliner, 1988). These beliefs are "non-negotiable," are associated to an object by individuals and are considered by them to as its definition. The second thesis showed that these beliefs also play a role in the dynamics of social representations, particularly when individuals adopt new behavior that contradicts them or makes previous behavior obsolete (Guimelli, 1988). A few years later, a third thesis was defended, this time at the University of Montpellier (Rateau, 1995). In this work, it was shown experimentally that the non-negotiable beliefs structuring representations are themselves hierachal. These works were our first contributions to social representation research, and apart from their theoretical implications, they also led to the finalization of specific methods dedicated to the study of social representations (Guimelli and Rouquette, 1992; Moliner, 1994; Rouquette and Rateau, 1998). Thus, they all served to answer the criticism of the SRT's first detractors.

But towards the end of the 1990s, new criticism appeared. At this time, it seemed as if social representation research was closing in upon itself, in utter disregard to

its obvious links with another up and coming trend; that of social cognition. For us, this criticism had to be taken into consideration, which is why we turned our research towards the systematic exploration of links between social representations and certain sociocognitive processes. All began with research on attitudes (Moliner and Tafani, 1997), followed by social categorization, attribution processes and social comparison processes (Rateau and Moliner, 2009). This work, mostly experimental, shows today that the barrier that some people saw between the social representation field and other areas of social psychology was probably just an illusion that time is beginning to erase. This is in any case our dearest wish as only this bridging of gaps will in the end allow us to fully [p. 488 ↓] understand and explain the social problems to which social psychology has the task of replying. It is also partly in this role that the SRT has been most successful as we now going to try and demonstrate.

The SRT's Applicability to Social Issues

To convince oneself of the SRT's applicability, one could try to list all the research that has adopted it. One would see in this case that numerous societal questions have been approached from this angle, and in fields as varied as health (e.g., Washer and Joffe, 2006), economy (e.g., Kirchler et al., 2003), marketing (e.g., Tafani et al., 2007), environmental psychology (e.g., Leone and Lesales, 2009), or relationships with new technology (e.g., Gal and Berente, 2008). However, apart from the fact that we are unable to make an exhaustive list of all these works in this chapter, it is not certain that such a list would allow the reader to understand why the SRT is used in such a diverse set of questions. From our point of view, the answer to this question depends on three points. The SRT is an adaptable and versatile theory, a common sense psychosocial theory, and is finally a theory that has given rise to the elaboration of varied methodologies.

A Flexible and Adaptable Theory

As we said earlier, one of the most frequent criticisms of SRT concerns the too-great imprecision of its concepts (McKinlay and Potter, 1987; Potter and Litton, 1985). And

it is true that upon reading Moscovici's original book, the apparent laxity with which the author presents the elements of his theory can be surprising, starting with the very definition he gives to the notion of social representations. But paradoxically, it is this very flexibility that confers on it its general scope. It is important to remember here that upstream from the theory, there is a protean phenomenon of which Durkheim had an intuition, and that Moscovici (2001a: 4) summarized with the words: "[T]he idea of social or collective representations is engraved in a societal vision in which coherency and practice are driven by beliefs, knowledge, norms and languages that it produces..." As such, it is a phenomenon that concerns logics of social relations just as much as those of action. And one whose, regulations can operate at different cognitive levels, including that of language. Thus, one understands the danger of attempting to study this kind of phenomena on the basis of concepts that are too narrow. This being the case, it is probably because the SRT's initial concepts are relatively broad that other disciplines, relatively unrelated to psychology, have been able to use them. Let us consider three examples.

The first is supplied by the work of historians who, wanting to transcend simple factual and event historiography, began to be interested by forms of thought and beliefs characteristic of past eras. Thus, they put the notion of "mentality" at the center of their preoccupations. Borrowed from Lévy-Bruhl (1922), this notion referred directly to that of mental representations, in relation to interactions in the social sphere. But it's clear today that the project of a "history of mentalities" comes down to a history of social representations.

The second example that we would like to briefly mention concerns geography. From the introduction of the mental map notion (Downs and Stea, 1977; Gould and White, 1974), and then the idea of a certain subjectivity in relation to space (Tuan, 1975, and finally the premise which recommends taking an interest in the mental processes which contribute to the perception of space, but which will especially lead to space being endowed with meanings and values. From this arises a "geography of representations," which considers representations to be finally determinants of spatial practice (Lussaut, 2007).

Finally, let us mention work carried out in linguistics, and more precisely in language didactics, where the necessity to understand [p. 489 ↓] the meanings associated with

learning and speaking a given language was noticed. This preoccupation has become central in multilingual situations, because of the identity problems they can give rise to. The notion of “linguistic representation” appeared (Dagenais and Jacquet, 2008), inspired directly by the SRT and designating beliefs relative to languages, their usages, and the groups that use them.

These examples suggest that outside the psychology field, when researchers ask themselves questions about cognitive determinants of behavior, they find in the SRT a conceptual framework that can be adapted to deal with their issues. But this is only possible thanks to the fact that this theory offers a great deal of latitude, which is, from our point of view, one of the reasons of its applicative success in social sciences.

An Everyday Knowledge Theory

Before being a belief or opinion theory, the SRT is first of all a theory of “common sense,” in that it accounts for the way in which common sense is formed, how it is structured, and how it combines with the preoccupations and social insertion of the people who use it. From this perspective, the most obvious application of the SRT concerns communication. Indeed, many studies show that different groups can have different representations of the same object. In fact, when these groups interact, whether it be for commercial reasons (a supplier and his clients), educational reasons (teachers versus pupils), or technical reasons (work teams), one can expect that different representations will be a potential source of confusion between groups. Consequently, the study of different existing representations can enable us to take suitable measures in the area of communication. For example, in his study on the representation of a hospital's computer security system, Vaast (2007) observed differences between doctors and nurses. For the doctors, system security meant principally access to data, whereas for nurses, it meant the protection of patient confidentiality. He concluded by insisting on the fact that the people responsible for the system have to take these differences on board in their personnel training.

Another SRT application is inspired by relations between representations and behavior. Indeed, generally speaking, common sense is what guides most of our every day behavior and interactions. “Our common sense includes a lot of know-how, ways

in which to make friends, succeed in life, and avoid crises, eat well, etc. ... It is on the basis of this knowledge that people are mostly aware of their situation or make important decisions..." (Moscovici, 2001b: 11). From this perspective, the study of social representations provides us with elements for understanding the reasons behind decisions or behavior. For example, in a study carried out on 1,005 French drivers, representations of speed were studied (Pianelli et al., 2007). This study showed that different representations coexist. The first one, the larger (44 percent of the population), was organized around the unique notion of "danger." Another one, the smaller (12 percent of the population) was organized around the sole notion of "pleasure." Thus, it was supposed that these two representations determined different driving practices. For those who belonged to the first representation, driving was seen as "careful," whereas the others saw it as "hedonistic." This hypothesis gains a first element of validation when one examines the causal link which the individuals made between speed and the occurrence of road accidents. Sixty-four percent of the "prudent" drivers thought that speed was the main cause of accidents, against only 24 percent of the "hedonists." Moreover, this study showed that there were less members of the first subgroup who admitted to having broken the speed limit than of the second subgroup (52 versus 76 percent on roads, and 47 versus 78 percent on motorways). There were also less people in the first subgroup to have been fined by the Police for speeding (9 versus 19 percent). As in many other [p. 490 ↓] studies the relationship between representations and behavior was clearly established. This relationship leads us to expect that action on the first will have an impact on the second. Thus, a third type of application appeared on the horizon, attempting to modify people's behavior. In fact, many studies (Mugny et al., 2000) show that influence procedures can provoke deep changes within social representations. But the few studies to have examined the durability of these changes have led to disappointing results.

However, recently, researchers working on these problems have been exploring a new avenue. It is no longer a question of modifying the contents of social representations, but of using these contents to bring individuals to make a decision. For example, Eyssartier et al. (2007) asked themselves the following question: How can the study of organ donation representations be useful for convincing people to become donors? From there, they identified four central elements and four peripheral elements of this representation. They then elaborated a "foot-in-the-door" technique (Freedman and

Fraser, 1966), designed to convince people to sign an organ donor card. It should be remembered that the foot-in-the-door principle consists of asking little (preparatory act) before asking more (final request). Yet, one knows that the importance that individuals grant to the preparatory act is a commitment-increasing factor (Kiesler, 1971, see also Burger, 1999). So the authors considered that a preparatory act referring to a central element of the representation was more important than a preparatory act referring to a peripheral element. Thus, they made the hypothesis that the effects of behavior commitment will be more effective in the first case than in the second. To test this hypothesis, an experimenter introduced himself as a volunteer for the “French Graft Establishment.” He only addressed people who were alone walking around a university hall, and asked them to sign a petition (preparatory act). This petition would allegedly be sent to the Ministry of Health, to gain financial aid for a communication campaign on organ donation. The petition's title contained either a slogan using a central element of the representation (i.e., “Organ donation helps others”), or a peripheral element of the representation (i.e., “Organ donation is a civic act”). Whether or not the preparatory act was accepted, the experimenter asked the person to sign an organ donor card (final request). Eight experimental conditions were studied (four “central slogans” and four “peripheral slogans”). The results showed that when the preparatory act concerned a central element of the representation of organ donation, there were significantly more participants who signed an organ donor card (51 percent) than when the preparatory act concerned peripheral elements (34 percent).

A Great Methodological Diversity

Without doubt, the SRT has provoked a remarkable diversity of methodologies because it can be applied to so many problems in various contexts. This methodological preoccupation became tangible from the end of the 1980s, when chapters dedicated to methodological questions were published in collective books on social representations. Later on, from the beginning of the 1990s, entire books were dedicated to methods of studying social representations.

The methodological advances presented in these books concern first of all techniques for collecting social representation content. Based on traditional psychosocial methods (interviewing, focus groups, investigations, etc.), verbal association techniques

emerged, aiming to minimize the amount of interpretation to be done by the researcher. To do this, these methods introduce constraints in the associative process, by inviting the subjects to only produce a certain type of answer (only verbs, adjectives or definitions). Furthermore, they invite the people questioned to evaluate their own contributions. For example, using the “Basic Cognitive Schema” technique (Guimelli, 1993, 1998), the participants are asked to say why they [p. 491 ↓] gave particular answers, and what inductors they used to do so. Using the Associative Network Method (De Rosa and Kirchler, 2001), the participants evaluate their associative production with help from different criteria supplied by the interviewer (positive or negative connotations, importance, etc.).

Methodological advances can also be found in questionnaire techniques. In contrast once again to more traditional approaches (opinion or attitude questionnaires), authors devise questionnaires that ask people to describe the studied objects in a standardized manner. It is no longer about measuring participants' opinions with regard to an object of representation, but rather highlighting the manner in which this object is described (see, for example, Moliner, 2002), and identifying the structuring elements of these descriptions.

Finally, the development of multivariate techniques, their computerization and their growth in accessibility have driven researchers to detail the specificities of each method compared with the SRT's postulates (see Doise et al., 1992).

Generally, researchers now have a large diversity of methods at their disposal, which helps them tackle a great range of questions in a large array of contexts.

Conclusion

By way of conclusion and introduction, we would like to further develop points that we have already mentioned rather allusively, because we think they constitute the basis of an important development in SRT, and more generally, of our knowledge of individual and group psychosocial functioning.

The first point refers to theoretical bridges that it seems possible to establish between the concepts of attitudes and social representation. This issue appears crucial and has already been the object of prolonged theoretical discussions (Billig, 1993; Farr, 1994; Howarth, 2006; Jaspar and Fraser, 1984; Scarbrough, 1990) in the attempt to understand and explain the reasons behind the mutual ignorance these two concepts have of each other. The inventory of these reasons, as interesting as it may be, would take too long to set out here. We prefer to focus on the hypothetical links that some authors have developed.

This is particularly true of Moliner and Tafani (1997) who consider that whatever theoretical definition is referred to, the observable part of attitudes always resides in the affective, behavioral, or cognitive responses that individuals express about an object. Yet, to produce this response, individuals need information about this object. This general idea has also been proposed by Fishbein and Ajzen (1975) and Zanna and Rempel (1988), for whom attitudes are the result of attributes that people associate with an object. In other, more direct words: to express an attitude towards an object, people have to have a representation of it.

This position is also defended by Doise (1989), for whom attitudes find their origin in more general knowledge of their social environment that individuals share. In studying this issue experimentally, Moliner and Tafani (1997) came to the conclusion that attitudes refer above all to evaluation, whereas representations refer above all to meaning. But to be able to evaluate an object, individuals have necessarily to have a meaning for it. In other words, attitudes are an evaluative expression of a shared representation of an object.

Rouquette (1996, 2010) also defends this idea, and recently proposed integrating the concepts of opinion, attitude, social representation and ideology in a global theoretical structure based on two general principles: the growing applicative stability and generality of each of these notions. From this dual viewpoint, Rouquette observes on the one hand that opinions are more volatile than attitudes (whence, for example, the need to do repeated opinion polls to measure fairly rapid fluctuations). On the other hand, he observes that opinions refer to particular objects, groups or individuals, in circumstances that are also particular, whereas attitudes, which are more [p. 492 ↓] general, refer to thematic categories involving more than one object. For example, an

opinion at a given time about a particular politician stems from the attitude towards politicians in general. In other words, a group's attitudes towards a given object are said to be the definitive source for opinions held about this object.

The same reasoning applies to the attitude/social representation duo. Apart from the first's larger variability than the latter's, it seems to be social representations that provide the basis for an attitude. Echebarria Echabe and Gonzalez Castro (1993) have shown, for example, that attitudes expressed by individuals towards elections are intimately intertwined with their representation of democracy in general.

Moreover, Rouquette proposes considering ideology as providing in its turn the basis for a social representation or a set of social representations. Certainly, ideology needs to be specified, because of its multiple meanings, its comprehensive scope, and its weakness of operationalization. But ideology can be conceived, not as a more or less organized assembly of content that may vary from one society to another, or from a group to its adversary, but as a repertory of general processes, with underlying formalizable qualities, and generic categories that are open to diverse description. It is essentially said to be made of values, norms, beliefs and themata (Moscovici and Vignaux, 2000). This framework obviously needs more development and formalization, but it is without doubt a promising basis for research with the aim of promoting a model of connection between the different conventions of expression of psychosocial functioning.

The second point refers to the links that can be made between the SRT and the processes traditionally studied in the field of social cognition which are the stereotypes, causal attribution or social comparison. Again, these two approaches have been strangers for a long time. The main issue, with out doubt, in this mutual ignorance concerns different perceptions of the "social." Social cognition advocates see social knowledge used by individuals as being the result of an accumulation of individual cognitive processes. Knowledge that is therefore, above all, individual, although shared. As for "social" determinisms, they are more than often limited to "others," thus totally neglecting laws, organizational structures, social relationships, or group history. As regards the advocates of social representation, it has for long time been considered that the processes described by social cognition were highly reductive, studied with the aid of methods that also appeared to be simplistic, and in the end totally incapable of

accounting for the historicity and impact of representations in the life of societies and in attitudes. But by wanting too much to account for this impact, studies devoted to social representations have often only led to a compilation of qualitative approaches, with blurred methodological contours, not allowing the restitution or the definition of the cognitive processes invested in their functioning.

However, we think it obvious to consider that the link between social cognition and social representations is twofold. On the one hand, we consider that social cognition processes intervene massively in the elaboration of social representations. It can be expected that the fruit of these processes (categories, stereotypes, causal attributions) are to be found in the contents and the structure of social representations. In other words, even if representations are collective constructions, they are still partially constructed by individuals.

At the same time, one can suppose that the processes studied in the field of social cognition are produced on the basis of representations. Thus, one can expect to observe modulations of these processes, depending on the underlying representations. To categorize, judge or explain one's immediate environment, individuals are thought to rely on, amongst other things, collective beliefs. This reflexive link is what unifies social representations and emotional, identity, attribution, social influence, or social comparison processes: social representations account for these [p. 493 ↓] processes, as well as actively participating in their own modes of operation. This idea can be illustrated by three examples.

First, in classic research, intergroup judgments and perceptions are studied through the processes of social categorization and stereotyping. But from the point of view of the SRT, intergroup representations are defined as social representations revolving around groups of people (Lorenzi-Cioldi and Clémence, 2001). Yet, a series of researches show that the central elements of an intergroup representation are the same as the stereotypical elements of the category of people they concern (Moliner and Vidal, 2003); that certain of these central elements play an explanatory role in the behaviors of members of the group in question (Moliner and Gutermann, 2004); and that they intervene to justify or rationalize asymmetric intergroup relations (Moliner et al., 2009).

Second, in the field of research performed on the process of attribution (Heider, 1958), the work of Ross (1977) highlighted the tendency of individuals to prefer dispositional factors (traits, aptitudes, motivation, etc.) to explain the behavior of an actor. On the other hand, we know that in situations of self-presentation, individuals prefer this type of explanation in order to give a good image of themselves, just as they judge more favorably people who prefer this type of explanation (Jellison and Green, 1981). However, in a series of experiments (Moliner, 2000), it was demonstrated that the expression of this preference remains dependent on the representations that the individuals activate in relation to the social situations in which they express themselves. Thus, when one suggests to the participants that the process of attribution to which they are going to submit themselves takes place in an affectively oriented social situation, one notices the disappearance of the systematic preference for dispositional explanations, in favor of the appearance of a self-serving bias (Zuckerman, 1979), or a person-positive bias (Sears, 1983). On the contrary, the systematic preference for dispositional factors is more marked when subjects make attributions in competitive situations with a practical purpose. Thus, the manner in which the subjects interpret the situation in which they find themselves at the moment when they are making the attributions determines the orientation of the process.

Third and finally, in the wide field of social comparison, several works have attested to the existence of a phenomenon of asymmetry in the comparison of the self to others (see Holyoak and Gordon, 1983; Mussweiller, 2001; Srull and Gaellick, 1983). The self and the other are seen as more similar when the other is taken as the point of departure (assimilation effect) and the reverse when it is the self that is taken as a reference point (contrast effect). In a series of recent studies (Chokier and Rateau, 2009; Rateau, submitted), we were able to demonstrate that this general process could be altered by the type of opinion at issue in the comparison and notably by the central or peripheral nature of the latter in structure of the representation of the object involved (in this case the social representation of studies shared by psychology students).

The participants are asked to compare themselves to a peer (either in the order the self-other, or other-self) who, depending on the case, is presented as defending a properipheral opinion, counterperipheral opinion, procentral, or countercentral in relation to studies. With regard to a peripheral opinion, characterized by a significant intragroup heterogeneity, the appearance of the “classic” process of interindividual comparison

of the self with another was recorded; that is to say, a contrast effect in the case of a comparison of the self–other order and an assimilation effect in the other–self order, regardless of the valency of the opinion defended by the source. This result illustrates perfectly the flexibility and possibilities for interindividual modulation which traditionally characterizes the peripheral elements of social representations.

With regard to a central opinion, the processes in play are very different. Whatever the order of the comparison, it is noticed that individuals differentiate themselves [p. 494 ↓] systematically from a group member who deviates from the central opinion and that they identify with a member who conforms. In other words, the individual seeks here to maintain the cohesiveness of the representation at any cost by systematically stating the “right” opinion in relation to the representation shared by their group of the object. The contrast/assimilation process does not depend here on the meaning of the comparison but only on the position taking displayed by the source, according to whether it contradicts the central opinion or not and ensures the homogeneity and social identity of the group.

This systematic study of the link between social representations and sociocognitive processes represents a desire to unite and mutually enrich both of these research fields. New hypotheses concerning functioning and roles of social representations, as well as sociocognitive processes when they are integrated into representational processes are beginning to appear. Let us wager that they will provide the basis of many studies, and that their theoretical and empirical range will be crucial in the development of our knowledge about the psychosocial functioning of individuals and groups.

Acknowledgments

The authors would like to sincerely thank Adam Chesterman and Anne Greaves for the translation of this chapter.

Patrick Rateau Pascal Moliner Jean-Claude Abric Pascal Moliner

References

- Abric, J.C. (1976) Jeux, conflits et représentations sociales . PhD thesis, Université de Provence, Aix-Marseille.
- Abric, J.C. Central system, peripheral system: their functions and roles in the dynamics of social representation . *Papers on Social Representations* , 2, 75–78. (1993)
- Abric, J.C. (2001) A structural approach to social representations . In K. Deaux, ed. and G. Philogène (eds), *Representations of the Social* , pp. 42–47. Oxford: Blackwell.
- Asch, S.E. Forming impressions of personality . *Journal of Abnormal and Social Psychology* , 41, 258–290. (1946) <http://dx.doi.org/10.1037/h0055756>
- Billig, M. (1993) Studying the thinking society: Social representations and attitudes . In G.M. Breakwell, ed. and D. Canter (eds), *Empirical Approach to Social Representations* , pp. 39–62. New York: Clarendon Press.
- Breakwell, G.M. and Canter, D.V. (1979) *Empirical Approach to Social Representations* . Oxford: Clarendon Press.
- Burger, J.M. The foot-in-the-door compliance procedure: A multiple-process analysis and review . *Personality and Social Psychology Review* , 3, 303–325. (1999) http://dx.doi.org/10.1207/s15327957pspr0304_2
- Clémence, A. (2001) Social positioning and social representations . In K. Deaux, ed. and G. Philogène (eds), *Representations of the Social* , pp. 83–95. Oxford: Blackwell.
- Codol, J.P. Influence de la représentation d'autrui sur l'activité des membres d'un groupe expérimental . *L'Année Psychologique* , 70, 131–150. (1970) <http://dx.doi.org/10.3406/psy.1970.27700>
- Cooper, J. and Croyle, R.T. Attitudes and attitude change . *Annual Review of Psychology* , 35, 395–426. (1984) <http://dx.doi.org/10.1146/annurev.ps.35.020184.002143>

Dagenais, D. and Jacquet, M. Theories of representation in French and English scholarship multilingualism . International Journal of Multilingualism , 5, 41–52. (2008)
<http://dx.doi.org/10.2167/ijm019.0>

Deaux, K. and Philogène, G. (2001a) Social Representations: Introductions and Explorations . New York: Blackwell.

Deaux, K. and Philogène, G. (2001b) Representations of the Social . Oxford: Blackwell.

De Rosa, A.S. and Kirchler, E. (2001) Ambiguous images in advertising: An application of the Associative Network Method . In C. Roland-Lévy, ed. , E. Kirchler, ed. , E. Penz, ed. and C. Gray (eds), Everyday Representations of the Economy , pp. 49–66. Wien: WUV Universitätsverlag.

Doise, W. (1989) Attitudes et représentations sociales . In D. Jodelet (ed.), Les Représentations Sociales , pp. 220–238. Paris: Presses Universitaires de France.

Doise, W., Clémence, A. and Lorenzi-Cioldi, F. (1992) Représenter les Représentations Sociales et Analyse de Données . Grenoble: Presses Universitaires de Grenoble;

J. Kaneko, ed. , tr. (1993) The Quantitative Analysis of Social Representations . Hemel-Hempstead: Harvester Wheatsheaf.

Downs, R.M. and Stea, D. (1977) Maps in Mind: Reflexions on Cognitive Mapping . New York: Harper and Row.

Durkheim, E. (1893) La Division du Travail . Paris: Presses Universitaires de France;

G. Simpson, ed. , tr. (1947) The Division of Labor in Society . New York: The Free Press.

Durkheim, E. (1895) Les Règles de la Méthode Sociologique . Paris: Alcan.

W.D. Halls, ed. , tr. (1982) The Rules of Sociological Method . New York: The Free Press.

Durkheim, E. Représentions individuelles et représentations collectives . Revue de Métaphysique et de Morale , 6, 273–302 (1898)

D.F. Pollock, ed. , tr. (1953) Individual and Collective Representations . in E. Durkheim (ed.) Sociology and Philosophy , pp. 1–34. London: Cohen and West.

Duveen, G. (2001) Representations, identities, resistance . In K. Deaux, ed. and G. Philogene (eds), Representations of the Social: Bridging Theoretical Tradition , pp. 257–270. Oxford: Blackwell.

Echebarria Echabe, A. and Gonzalez Castro, J.L. Social representation of power and democracy, attitudes towards elections and voting behaviour . International Review of Social Psychology , 6, 21–46. (1993)

Eyssartier, C., Joule, R.V. and Guimelli, C. Effets comportementaux et cognitifs de l'engagement dans un acte préparatoire activant un élément central vs périphérique de la représentation du don d'organe . Psychologie Française , 52, 499–517. (2007)

Farr, R.M. Attitudes, social representations and social attitudes . Papers on Social Representations , 3, 33–36. (1994)

Farr, R. and Moscovici, S. (1984) Social Representations . Cambridge: Cambridge University Press.

Fishbein, M. and Ajzen, I. (1975) Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research . Boston: Addison-Wesley.

Flament, C. Image des relations amicales dans les groupes hiérarchisés . L'Année Psychologique , 71, 117–125. (1971) <http://dx.doi.org/10.3406/psy.1971.27725>

Flament, C. (1989) Structure et dynamique des représentations sociales . In D. Jodelet (ed.), Les Représentions Sociales , pp. 204–219. Paris: Presses Universitaires de France.

Freedman, J.L. and Fraser, S.C. Compliance without pressure: The foot-in-the-door techniques . *Journal of Personality and Social Psychology* , 4, 195–202. (1966) <http://dx.doi.org/10.1037/h0023552>

Freud, S. (1908) *On the Sexual Theories of Children*, Vol. 9 . London: Standard Edition.

Freud, S. (1922) *Some Points in a Comparative Study of Organic and Hysterical Crisis*, Vol. 1 . London: Hogarth Press.

Gal, U. and Berente, N. A social representation perspective on informational implementation: rethinking the concept of frames . *Information, Technology and People* , 21, 133–154. (2008) <http://dx.doi.org/10.1108/09593840810881051>

Gould, P. and White, R. (1974) *Mental Maps* . Harmondsworth: Penguin Books. <http://dx.doi.org/10.4324/9780203163757>

Guimelli, C. (1988) Aggression idéologique, pratiques nouvelles et transformation progressive d'une représentation sociale . PhD thesis, Université de Provence, Aix en Provence.

Guimelli, C. Locating the central core of social representations: towards a method . *European Journal of Social Psychology* , 23, 555–559. (1993) <http://dx.doi.org/10.1002/ejsp.2420230511>

Guimelli, C. Differentiation between the central core elements of social representations: normative vs. functional elements . *Swiss Journal of Psychology* , 57, 209–224. (1998)

Guimelli, C. and Rouquette, M.L. Contribution du modèle associatif des Schèmes Cognitifs de Base à l'analyse structurale des représentations sociales . *Bulletin de Psychologie* , 405, 196–202. (1992)

Heider, F. (1958) *The Psychology of Interpersonal Relations* . New York: Wiley. <http://dx.doi.org/10.1037/10628-000>

Holyoak, K.J. and Gordon, P.C. Social reference points . *Journal of Personality and Social Psychology* , 44, 881–887. (1983) <http://dx.doi.org/10.1037/0022-3514.44.5.881>

Howarth, C. How social representations of attitudes have informed attitudes theories: the consensual and the reified . *Theory and Psychology* , 16, 691–714. (2006) <http://dx.doi.org/10.1177/0959354306067443>

Jaspar, J. and Fraser, C. (1984) Attitudes and social representations . In R.M. Farr, ed. and S. Moscovici (eds), *Social Representations* , pp. 101–123. Cambridge: Cambridge University Press.

Jellison, J. and Green, J. A self presentation approach to the fundamental attribution error: the norm of internality . *Journal of Personality and Social Psychology* , 40, 643–649. (1981) <http://dx.doi.org/10.1037/0022-3514.40.4.643>

Jodelet, D. (1989) *Folie et Représentations Sociales* . Paris: Presses Universitaires de France;

T. Pownall, ed. , tr. (1992) *Madness and Social Representations* . Berkeley: University of California Press.

Kiesler, C.A. (1971) *The Psychology of Commitment. Experiments Linking Behavior to Belief* . New York: Academic Press.

Kirchler, E., Maciejovsky, B. and Schneider, F. Everyday representations of tax avoidance, tax evasion, and tax flight: Do legal differences matter? *Journal of Economic Psychology* , 24, 535–553. (2003) <http://dx.doi.org/10.1016/S0167-4870%2802%2900164-2>

Kronberger, N. and Wagner, W. (2000) Key words in context: statistical analysis of text features . In M.W. Bauer, ed. and G. Gaskell (eds), *Qualitative Researching with Text, Image and Sound. A Practical Handbook* , pp. 299–317. London: Sage.

Kruglanski, A.W. (2001) Social cognition, social representations and the dilemmas of social theory construction . In K. Deaux, ed. , and G. Philogène (eds), *Representations of the Social* , pp. 242–248. Oxford: Blackwell.

Leone, F. and Lesales, T. The interest of cartography for a better perception and management of volcanic risk: From scientific to social representations .

Journal of Volcanology and Geothermal Research , 186, 186–194. (2009) <http://dx.doi.org/10.1016/j.jvolgeores.2008.12.020>

Lévi-Strauss, C. (1962) *La pensée sauvage* . Paris: Plon;

G. Weidenfeld, ed. and Nicolson, ed. , tr. (1966) *The Savage Mind* . Chicago: University of Chicago Press

Lévy-Bruhl, L. (1922) *La mentalité primitive* . Paris: Alcan;

L.A. Clare, ed. , tr. (1978) *Primitive Mentality* . New York: AMS Press.

Linton, R. (1945) *The Cultural Background of Personality* . New York: D. Appleton-Century Co.

Lorenzi-Cioldi, F. and Clémence, A. (2001) Group processes and the construction of social representations . In M.A. Hogg, ed. and S. Tindale (eds), *Handbook of Social Psychology: Group Processes* , pp. 311–333. Oxford: Blackwell.

Lorenzi-Cioldi, F. and Clémence, A. (2010) Social representations . In J. Levine, ed. and M. Hogg (eds), *Encyclopedia of Group processes and Intergroup Relations* . London: Sage. <http://dx.doi.org/10.4135/9781412972017>

Lussaut, M. (2007) *L'homme Spatial. Construction Sociale de l'espace Humain* . Paris: Seuil.

Markova, I. Language and authenticity . *The Journal for the Theory of Social Behaviour* , 27, 265–275. (1997) <http://dx.doi.org/10.1111/1468-5914.00037>

Markova, I. (2003) Dialogicality and Social Representations. *The Dynamics of Mind* . Cambridge: Cambridge University Press.

Mauss, M. (1903) *Représentations Collectives et Diversité des Civilisations* . Paris: Minuit (1974 edition).

McKinlay, A. and Potter, J. Social representation: A conceptual critique . Journal for the Theory of Social Behavior , 17, 471–487. (1987) <http://dx.doi.org/10.1111/j.1468-5914.1987.tb00109.x>

Moliner, P. (1988) La représentation sociale comme grille de lecture . PhD thesis, Université de Provence, Aix-en-Provenc.

Moliner, P. (1994) Les méthodes de repérage et d'identification du noyau des représentations sociales . In C. Guimelli (ed.), Structures et Transformations des Représentations Sociales , pp. 199–232. Neuchâtel: Delachaux et Niestlé.

Moliner, P. De la norme d'internalité à la représentation des relations sociales . Revue Internationale de Psychologie Sociale , 2, 7–32. (2000)

Moliner, P. Ambiguous scenario and attribute challenge techniques. Social Representations of 'The Firm' and 'The nurse' . European Review of Applied Psychology , 3, 273–280. (2002)

Moliner, P. and Gutermann, M. Dynamique des descriptions et des explications dans une représentation sociale . Papers On Social Representations , 13, 1–12. (2004)

Moliner, P., Lorenzi-Cioldi, F. and Vinet, E. Utilité sociale des représentations intergroupes de sexe. Domination masculine, contexte professionnel et discrimination positive . Cahiers Internationaux de Psychologie Sociale , 83, 25–44. (2009)

Moliner, P. and Tafani, E. Attitudes and social representations: A theoretical and experimental approach . European Journal of Social Psychology , 27, 687–702. (1997) <http://dx.doi.org/10.1002/%28SICI%291099-0992%28199711/12%2927:6%3C687::AID-EJSP839%3E3.0.CO;2-7>

Moliner, P. and Vidal, J. Stéréotype de la catégorisation et noyau de la représentation . Revue Internationale de Psychologie Sociale , 1, 157–176. (2003)

Moscovici, S. (1961). La Psychanalyse: Son Image et son Public . Paris: Presses Universitaires de France (1976 edition);

D. Macey, ed. , tr. (2008) Psychoanalysis. Its Image and Its Public . Cambridge: Polity Press.

Moscovici, S. (1969) Préface . In C.Herzlich, Santé et Maladie , pp. 7–12. Paris: Mouton.

Moscovici, S. (1981) On social representations . In J. Forgas (ed.), Social Cognition: Perspectives on Everyday Understanding , pp. 181–210. New York: Academic Press.

Moscovici, S. (1982) The coming era of representations . In J.P. Codol, ed. and J.P. Leyens (eds), Cognitive Approaches to Social Behavior , pp. 115–150. The Hague: M. Nijhoff.

Moscovici, S. (1984) The phenomenon of social representations representations . In R. Farr, ed. and S. Moscovici (eds), Social Representations , pp. 3–69. Cambridge: Cambridge University Press.

Moscovici, S. Notes towards a description of social representations . European Journal of Social Psychology , 18 (3) 211–250. (1988) <http://dx.doi.org/10.1002/ejsp.2420180303>

Moscovici, S. (1989) Des représentations collectives aux représentations sociales. Eléments pour une histoire . In D. Jodelet (ed.), Les Représentations Sociales , pp. 62–86. Paris: Presses Universitaires de France.

Moscovici, S. (2001a) Social Representations: Explorations in Social Psychology . Cambridge: Polity Press.

Moscovici, S. (2001b) Why a theory of social representations? In K. Deaux, ed. , and G. Philogène (eds), Representations of the Social , pp. 8–36. Oxford: Blackwell.

Moscovici, S. and Vignaux, G. (2000) The concept of themata . In S. Moscovici, ed. and G. Duveen (eds), Social Representations. Explorations in Social Psychology , pp. 156–183. Cambridge: Polity Press.

Mugny, G. and Carugati, F. (1989) Social Representations of Intelligence . Cambridge: Cambridge University Press.

Mugny, G., Tafani, E., Falomir, J.M. and Layat, C. Source credibility, social comparison and social influence . International Review of Social Psychology , 13, 151–175. (2000)

Mussweiller, T. Focus of comparison as a determinant of assimilation versus contrast in social comparison . Personality and Social Psychology Bulletin , 27, 38–47. (2001)
<http://dx.doi.org/10.1177/0146167201271004>

Piaget, J. (1932) Le Jugement Moral Chez l'Enfant . Paris: Presses Universitaires de France;

M. Gabain, ed. , tr. (1965) The Moral Judgment of the Child . New York: The Free Press.

Pianelli, C., Saad, F. and Abric, J.C. (2007) Social representations and acceptability of LAVIA (French ISA system) . Proceedings of the 14th World Congress and Exhibition on Intelligent Transport Systems and Services , 8–12 October 2007, Pékin, Paper 4057.

Potter, J. and Litton, I. Some problems underlying the theory of social representations . British Journal of Social Psychology , 24, 81–90. (1985) <http://dx.doi.org/10.1111/j.2044-8309.1985.tb00664.x>

Rateau, P. (1995) Structure et fonctionnement du système central des représentations sociales . PhD thesis, Université de Montpellier 3, Montpellier.

Rateau, P. (submitted) Assimilation and contrast to ingroup prime: The moderating effect of central vs. peripheral elements of social representations . International Review of Social Psychology.

Rateau, P. and Moliner, P. (2009) Représentions Sociales et Processus Sociocognitifs . Rennes: Presses Universitaires de Rennes.

Ross, L. (1977) The intuitive psychologist and his shortcoming: Distortions in the attribution process . In L. Berkowitz (ed.), Advances in Experimental Social

Psychology , 10, 173–220. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960357-3>

Rouquette, M.L. (1996) Représentions et idéologie . In J.C. Deschamps, ed. and J.L. Beauvois (eds), Des Attitudes aux Attributions , pp. 163–173. Grenoble: Presses Universitaires de Grenoble.

Rouquette, M.L. (2010) La Pensée Sociale . Toulouse: Erès.

Rouquette, M.L. and Rateau, P. (1998) Introduction à l'Étude des Représentions Sociales . Grenoble: Presses Universitaires de Grenoble.

Scarborough, E. (1990) Attitudes, social representations and ideology . In C. Fraser, ed. and G. Gaskell (eds), The Social Psychological Study of Widespread Beliefs , pp. 99–117. New York: Clarendon Press.

Schank, R. and Abelson, R. (1977) Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures . Hillsdale: Erlbaum.

Sears, D.O. The person-positive bias . Journal of Personality and Social Psychology , 44, 223–250. (1983) <http://dx.doi.org/10.1037/0022-3514.44.2.233>

Simmel, G. (1908) Soziologie . Berlin: Duncker Hamblot;

S. Muller, ed. and L. Deroche-Gurcel, ed. , tr. French (1999) Sociologie. Etudes sur les Formes de la Socialisation . Paris: Presses Universitaires de France.

Smith, A. (1776) An Inquiry into the Nature and Causes of the Wealth of Nations . London: Strahan & Cadell.

Spini, D. Multidimensional scaling. A technique for the analysis of the common field of social representations . European Review of Applied Psychology , 52, 231–240. (2002)

Srull, T.K. and Gaelick, L. General principles and individual differences in the self as a habitual reference point: An examination of self-other judgments of similarity . Social Cognition , 2, 108–121. (1983) <http://dx.doi.org/10.1521/soco.1983.2.2.108>

Tafani, E., Haguel, V. and Menager, A. From corporate images to social representations: an application to the motor car industry . *Les Cahiers Internationaux de Psychologie Sociale* , 73, 27–46. (2007)

Tuan, Y-F. Images and mental maps . *Annals of the Association of American Geographers* , 65, 205–212. (1975) <http://dx.doi.org/10.1111/j.1467-8306.1975.tb01031.x>

Vaast, E. Danger is in the eye of the beholders: Social representations of information systems security in healthcare . *Journal of Strategic Information Systems* , 16, 130–152. (2007) <http://dx.doi.org/10.1016/j.jsis.2007.05.003>

Vergès, P. (1996) Bibliographie des représentations sociales . Third International Conference on Social Representations, Aix-en-Provence, September 27–30.

Wagner, W. Fields of research and sociogenesis of social representations: a discussion of criteria and diagnostics . *Social Science Information* , 33, 199–228. (1994) <http://dx.doi.org/10.1177/053901894033002004>

Wagner, W., Duveen, G., Farr, R., Jovchelovitch, S., Lorenzi-Cioldi, F., Markovà, I. and Rose, D. Theory and method of social representations . *Asian Journal of Social Psychology* , 2, 95–125. (1999) <http://dx.doi.org/10.1111/1467-839X.00028>

Washer, P. and Joffe, H. The ‘hospital superbug’: Social representations of MRSA . *Social Science and Medicine* , 63, 2141–2152. (2006) <http://dx.doi.org/10.1016/j.socscimed.2006.05.018>

Weber, M. (1921) *Economy and Society* . Berkeley, CA: University of California Press, 1978.

Zanna, M. and Rempel, J. (1988) Attitudes: a new look at an old concept . In D. Bar-tal, ed. and A.W. Kruglanski (eds), *The Social Psychology of Knowledge* , pp. 315–334. New York: Cambridge University Press.

Zuckerman, M. Attribution of success and failure revisited or: The motivational bias is alive and well in attribution theory . *Journal of Personality and Social Psychology* , 47, 245–287. (1979)

<http://dx.doi.org/10.4135/9781446249222.n50>

Handbook of Theories of Social Psychology

A Theory of Individualism and Collectivism

Contributors: Harry C. Triandis & Michele J. Gelfand

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology

Chapter Title: "A Theory of Individualism and Collectivism"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249222

DOI: <http://dx.doi.org/10.4135/9781446249222.n51>

Print pages: 498-521

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249222.n51>

[p. 498 ↓]

Chapter 51: A Theory of Individualism and Collectivism

Abstract

The evolution of individualism and collectivism theory and research is reviewed. The antecedents of collectivism–individualism can be found in the ecology, family structure, wealth distribution, demography, history, cultural diffusion, and situational conditions. The consequences of collectivism–individualism include differences in attention, attribution, cognition, emotion, motivation, self-definitions, values, language use, and communication, as well as other kinds of social and organizational behavior. Applications of individualism and collectivism include improvements in conflict resolution, health, international relations, and cross-cultural training.

The Evolution of Individualism and Collectivism Theory and Research

Culture is to society what memory is to individuals (Kluckhohn, 1954). It consists of what “has worked” in the experience of a group of people so it was worth transmitting to peers and descendants. Another definition of culture was provided by anthropologist Redfield (1941): “Culture is shared understandings made manifest in act and artifact.” In short, culture is shared behavior and shared human-made aspects of the society. Thus, it includes “practices” (the way things are done here) and “values” (the way things should be done). These older definitions of culture focus on what is outside the person (e.g., do people drive to the right or left). The more recent definitions also stress what is inside the person (e.g., is the self independent or interdependent of in-groups). Almost every aspect of psychological functioning is influenced, to some extent, by culture.

Thus, it is best to view culture and psychology as making each other up (Cole, 1996; Schweder, 1990).

Cultures differ in a myriad of ways. By far, the most well-researched dimension of culture to date is *individualism* and *collectivism*. [p. 499 ↓] Within the twentieth century, there has been extensive discussion of the constructs in sociology (e.g., Durkheim, 1933; Parsons, 1949; Riesman et al., 1961), anthropology (Kluckhohn, 1956; Mead, 1967; Redfield, 1956), and psychology (Chinese Culture Connection, 1987; Hofstede, 1980; Schwartz, 1994; Triandis, 1995). Across each of these disciplines, scholars have been concerned with the nature of the relationship between the individual and the group. This theme has also been referred to as *self-emphasis* and *collectivity* (Parsons, 1949), *Gesellschaft* and *Gemeinschaft* (Toennies, 1957), *mechanical* and *organic solidarity* (Durkheim, 1933), *individualism* and *collaterality* (Kluckhohn and Strodtbeck, 1961), *agency* and *community* (Bakan, 1996), *individualism* and *collectivism* (Hofstede, 1980), *autonomy* and *conservation* (Schwartz, 1990), and *independence* and *interdependence* (Markus and Kitayama, 1991). Although there are subtle differences in meanings of these terms, they all relate to a theme which contrasts the extent to which people are autonomous individuals or embedded in their groups (Hofstede, 1980; Schwartz, 1994; Triandis, 1989).

Regardless of their labels, these constructs have been at the cornerstone of theory and research in culture and psychology. In this chapter, we trace the evolution of theory and research on individualism and collectivism, discussing their role first in ancient legal and religious institutions, in later political theory in the nineteenth century, and in empirical work in psychology in the mid and latter twentieth century. Although individualism–collectivism cannot be described as unified theory per se, research over the last four decades has illuminated the defining features of the constructs, their ecological, situational, and dispositional antecedents, and a wide range of consequences that they have for social psychological and organizational phenomena. To be sure, our review in this chapter is necessarily selective, as research in this area is extensive and warrants a volume into itself. After describing the evolution of theory and research on the constructs, we provide an evaluation of the constructs and discuss practical implications that have been derived from this collective research effort.

History of the Individualism and Collectivism Constructs: Contributions from Law, Religion, and Philosophy

Although individualism and collectivism have been by now well researched in psychology, discussions of the constructs can be found in ancient legal writings, religious texts, and moral-political philosophies. The contrasts between the individualism and collectivism constructs were first found in ancient legal structures in the Middle East in the laws of Hammurabi. In particular, the King of Babylonia (1792–1750 bc) is credited with establishing some of the world's first written laws, wherein universal codes of behaviors supplanted a focus on the rights of individuals. At the time, the codes replaced the more individualistic notion of tit-for-tat retaliation (revenge) with a system of monetary fines that were applied universally. More generally, the code of Hammurabi identified the need for individuals to maintain positive relations with others, lest they face heavy sanctions. The recognition of individuals as being interdependent and having duties and obligations to other group members are defining attributes of the cultural construct that we now call collectivism (Triandis, 1995). Notably, the Code of Hammurabi was not the only legal expression of the collectivist cultural construct in the ancient Middle East. Codes of conduct that centered on creating group standards of behavior were also part of the law of the Hebrews in the book of the Law of Moses (Kagan, 1966), the purpose of which was also to establish standards for individual behavior to protect the group, rather than to allow individual preferences to determine what is right and what is wrong (Durant, 1935).

[p. 500 ↓]

As cultures later developed, the more individualistic notion of rational principles and individual rights became more prevalent within legal systems. The practice of presenting individual cases before formally appointed judges was prevalent in Athens and in Rome. For example, in the Law of Cincius (204 bc), legal representation was viewed as the most effective way to present the facts of a particular case. Importantly, this system was seen as superior to interpretations of right and wrong based on codes of normative

behavior alone, the latter of which was common in earlier centuries and was more akin with collectivism.

The constructs of individualism and collectivism were also manifested in religious institutions throughout the centuries. In the West, concerns with group identity and ingroup–outgroup distinctions, both attributes of collectivism, can be seen in religious philosophies and practices. The ancient Hebrews' religion was based on the strong ethnic identity of the Jews (Durant, 1935), and was predicated on the belief that the group was the “chosen one” of God, as compared with other groups. Other religious groups also viewed their religions as a form of group identity as contrasted to other groups. For example, in the Koran of the Moslems it is stated – “Believers, take neither the Jews nor the Christians for your friends” (Dawood, 1956). Likewise, within the Christian tradition, for individuals to be saved they had to embrace the Christian God as the only one true God and reject other conceptions of god that were found in other religions.

Religions in the East were much more focused on duties and obligations within a hierarchical structure, which is associated with some forms of modern-day collectivism. In India, and in ancient Japan, caste systems were also developed, and group identity was even further reinforced within a legal system that held entire families responsible for individual members' actions (Durant, 1935). These notions of group accountability predated empirical work in psychology which centuries later has indeed shown that East Asians hold many people, particularly groups and their leaders, accountable for a given action (Chiu and Hong, 1992; Chiu et al., 2000; Menon et al., 1999; Zemba et al., 2006).

Likewise, in China, Confucian philosophy emphasized the importance of group identity, conformity, and long-term relationships. Confucius also stressed the importance of obligations that individuals have within their family, within the nation, and within the world at large. For example, Confucian philosophy dictates that individuals are required to respect their fathers and elder brothers so as to maintain family harmony. This prepared the individual to respect the structures of the state, which was needed to maintain national harmony. Throughout his writings, Confucius emphasized the importance of subjugating personal wants and desires for the greater good of the group (Streep, 1995). This philosophy, while dating back 4,000 years, is still prevalent in much of Eastern Asia today.

The nature of the relationship of the individual to the state was also at the center of much philosophical thought and debate in the late eighteenth century. Conceptions of individualism were associated with liberalism and included the ideas of maximum freedom of the individual, voluntary groups that individuals can join or leave, and equal participation of individuals in group activities (*Encyclopedia Britannica*, 1953, Vol. 12: 256a). As a moral-political philosophy, liberalism placed a great importance on the freedom of individuals to use reason to make personal choices, and to have rights to protect these freedoms (Kim, 1994). Across societies, the importance of the freedom of individuals was also reflected in the American Revolution (all humans are created equal, and pursuit of happiness is their fundamental right) and the French Revolution (liberty, equality, fraternity).

At the same time, other philosophers, most notably Jean Jacques Rousseau, emphasized the importance of the collective over the individual. In the *Social Contract* (1762), Rousseau argued that the individual is only [p. 501 ↓] free by submitting to the general will. The general will was conceived as the common core of opinion that remains after private wills cancel each other out. Rousseau argued that the general will, which can be ascertained by majority voting, is “always right and tends to the public advantage” (*Encyclopedia Britannica*, 1953, Vol. 12: 256a).

Later, within the nineteenth century, the French intellectual Alexis de Tocqueville elaborated upon the concept of individualism based on his travels throughout the United States (Bellah et al., 1985). De Tocqueville used the term “individualism” in connection with democracy in American society and contrasted the American social structure with those found in the aristocratic European tradition. Later, political philosophers such as Dewey (1930), Dumont (1986), and Kateb (1992) also discussed ideas related to individualism. Dewey (1930) distinguished what he referred to as “old” individualism, which included the liberation from legal and religious restrictions, from the “new” individualism, which focused on self-cultivation. Dumont (1986) argued that individualism was a consequence of Protestantism (i.e., humans do not have to go to church to communicate with God), political developments (emphasis on equality and liberty), and economic developments (e.g., affluence).

In all, the constructs of individualism and collectivism received much theoretical attention in legal, religious, and philosophical writings for centuries. It wasn't until the

1960s, however, that they began to receive systematic empirical attention in the field of psychology.

Individualism and Collectivism: Contributions from Psychologists across Four Decades

The development of theory and research on individualism and collectivism in psychology truly represents a collective effort of many psychologists. Rather than being primarily driven by one source of influence, there have been numerous people who have made important contributions. As the ancient Hindu saying dictates, “Truth is one; it has many names;” so too is the case in terms of individualism and collectivism theory and research. In what follows we trace the evolution of theories and research and the serendipity of collaborations and discoveries that have shaped and continue to shape our knowledge of the constructs.

On the Origins of Individualism– Collectivism Theory and Research: The Analysis of Subjective Culture

Arguably the first empirical evidence for individualism and collectivism in psychology can be traced to a large multination project known as the *Analysis of Subjective Culture* (Triandis, 1972). The concept of culture, at that time, was reflected in the work of three anthropologists: Clyde Kluckhohn (1954), Melvin Herskovits (1955), and Ralph Redfield (1941). Herskovits had defined culture as the human-made part of the environment. The human-made part of the environment consists of *physical* (e.g., tools, bridges, educational systems, religious institutions), as well as *subjective* elements (e.g., beliefs, attitudes, norms, values). Following this lead, in the *Analysis of Subjective Culture*, Triandis (1972) set out to develop a theory and psychological methods to study

“subjective culture.” At the time, there were no psychological methods available and no overarching framework that examined the psychological underpinnings of culture.

The project that resulted in this publication, like many others described in this book, was the result of serendipitous events. In 1963, the Chief of Naval Operations asked the Office of Naval Research (ONR) to organize and support research so to better prepare Naval officers to manage cultural differences. ONR came to the psychology [p. 502 ↓] department of the University of Illinois and asked us if it could undertake this project. Fred Fiedler put together a team that consisted of Charles Osgood, Larry Stolurow, and Harry Triandis. Triandis already had some crosscultural experience, so it was natural that he was given the job of analyzing “culture” so that it could be converted into computer-supported teaching experiences by Stolurow, and communicated to seamen (Osgood), who would be organized into teams whose leadership would be studied by Fiedler.

Triandis was already familiar with culture-specific (*emic*) and culture-general (*etic*) constructs (Triandis, 1964), so he aimed at developing methods that could include some *etic* elements, so that cultures could be compared, as well as *emic* elements from each culture so that one could understand each culture from the “inside.”¹ Many paper-and-pencil methods were developed which tapped different elements of “subjective culture” including, categorizations, associations, attitudes, beliefs, expectations, and norms, among others. What became clear in the analysis was that *coherent themes* cut across these different elements of subjective culture. The theme of *individualism and collectivism* began to emerge from psychological data for the first time across numerous countries, including the US, Greece, India, and Japan.

Elements of Subjective Culture

The basic element for the study of culture is *categorization*. What stimuli are treated as equivalent by members of the culture? Members of each culture have unique ways of categorizing experience. For example, who is a member of my ingroup? In some cultures it is those who were born in the same place, or belonged to the same tribe, race, social class, religion, or who were blood relatives. In other cultures it is people

who “think like I do.” The two categories might overlap, but they are not the same. Thus, there are both *etic* (common elements: my group) and *emic* elements (culture-specific elements: specific groups) defining this category. Later work (Brewer and Yuki, 2007) distinguished two kinds of definitions of ingroup: a *category-based identity*, such as “I am an American” and a *relationship-based identity*. In all, by studying how people categorize experience we learn much about their culture.

Members of each culture have unique ways of *associating* one category with another. For example, is “socialism” referring to a political party, an ideology, or both? Are “fathers” in this culture assumed to be severe or lenient, with respect to children of different ages? In addition, cultures differ in the kinds of perceived antecedent-consequent relationships that people use (e.g., if you have “hard work” then you have “progress;” if you have “progress” then you have “health”), *attitudes* (e.g., is “socialism” good or bad, would members of the culture support a socialist party?), *beliefs* (e.g., “socialism” results in good health; or results in an impoverished society), *expectations* (e.g., if there is socialism then there is poverty), *ideals* (e.g., widows should not be passionate), *memories* (e.g., I remember the names of each of my cows), *norms* (e.g., members of this society give their seat to old people), *role perceptions* (e.g., the mother–son role is warmer than the father–son role in this culture), *stereotypes* (e.g., lower class people are not intelligent), *tasks* (e.g., to make this tool one has to first get some redwood), *values* (e.g., “security” is very important). Later work by Triandis (1977, 1980) resulted in a model linking *behavioral intentions* (e.g., I intend to do X) and *behavior* (X), which included also *norms* (most people I respect think I should do X), *self-definitions* (e.g., I am the kind of person who does X), *habits* (e.g., I frequently do X), and *facilitating conditions* (e.g., I am highly aroused to do X, I am capable of doing X; the situation calls for me to do X) The theory of reasoned action (Fishbein and Ajzen, 1975) was combined with elements of the 1980 model to form the unified theory of behavior (Fishbein et al., 2001). In addition, some “probes” (Triandis, 1972) into subjective culture were carried out, by studying stereotypes, [p. 503 ↓] antecedent-consequent relationships, which also provided clues about the values of each culture, and role perceptions.

Cultural Syndromes: Interrelated Elements of Subjective Culture

Most importantly, Triandis (1972) examined if the information obtained from each of the “probes” into culture, by the various methods, hangs together. Indeed, a comparison of multiple measures across the US and Greece, showed that the data from each of the methods can be summarized by using certain themes. Specifically, the Greek data indicated much more contrast between behaviors toward the ingroup and outgroup members in Greece than in the US. This turned out to be a major characteristic of collectivist cultures (Triandis, 1995). The traditional Greeks (of the 1960s, when the data were collected) defined their universe in terms of the triumphs of their ingroup over their outgroups, while for Americans this worldview was of little or no importance. In Greece, relationships with authorities and social relations in general reflected the ingroup–outgroup relationships where there is much association and intimacy and low hostility within ingroups whereas in the US, participants express some hostility within ingroups and emotional distance from ingroup members. An ingroup in Greece was defined as a group of individuals about whose welfare a person is concerned, with whom the person is willing to cooperate without demanding equitable returns, and separation from whom leads to anxiety.

Foreshadowing the large literature on culture and self, Triandis (1972) found that Greek self-definition depended on the way ingroup members saw the person, thus individuals' worth was defined by the group. By contrast, American self-definition depended on the way individuals saw themselves. Consistent with the collectivist–individualist contrast, the concept MYSELF was rated (on semantic differential scales) “stronger” by Americans than by Greeks, but the concept MY RELATIVES was rated stronger by Greeks than by Americans. Greeks also perceived behaviors in *context* to a greater extent than did Americans, an attribute that later proved to be a key characteristic of collectivist cultures. For example, CHEATING was completely unacceptable for the Greeks when the target was an ingroup member, but was perfectly okay if the target was an outgroup member.

Ecocultural Framework of Dimensions of Culture

Another contribution of the *Analysis of Subjective Culture* was that it placed the thematic elements of subjective culture into a larger ecological and historical framework. The theoretical framework that was developed included *distal antecedents* (e.g., climate) and *historical events* (e.g., wars), *proximal antecedents* (e.g., occupations, language used, religion), and *immediate antecedents of action* (which included all the elements listed in the following paragraphs), which result in patterns of action. For example, Greece is cut up into small segments, because of the numerous mountains and islands, and that results in ingroups that are linked to place. The 350-year Ottoman occupation required knowing who could be trusted (i.e., who was ingroup). Furthermore, competition for scarce resources made it difficult to be cooperative with outgroups. The framework offered other ecological antecedents of subjective culture. For example, when resources are abundant there is more individualism. Cultures that are relatively isolated from other cultures, and in which making a living requires people to work together very frequently are likely to be more collectivistic. Climate is also an important factor in shaping subjective culture. For example, self-expression is higher in wealthy countries with harsh climates (cold or hot) than in countries with temperate climates, whereas self-expression is lower in poor [p. 504 ↓] countries with harsh climates than in poor countries in temperate climates (Van de Vliert, 2007).

In all, the *Analysis of Subjective Culture* was the first systematic study that illustrated cross-cultural differences in the emphasis on individuals versus groups, and began to trace the ecological and historical correlates of the constructs. As reviewed below, this became a central focus in later work in cross-cultural social psychology.

Development of the Individualism–Collectivism Constructs: The 1980s

At the same time as Triandis was finalizing the *Analysis of Subjective Culture* findings, he happened to meet Geert Hofstede who was also collecting data on the constructs. It was 1971, at the Congress of the International Association of Applied Psychology, in Liege, Belgium, that Hofstede mentioned to Triandis that the dataset existed. Hofstede took Triandis to his office in Brussels and they there discussed the analysis of the dataset. A factor analysis included a factor which Hofstede named *individualism–collectivism* – arguably the first formal use of the terms in psychology. The factor had a strong similarity to the American–Greek dataset of Triandis (1972), and Hofstede (1980) later referred to Triandis (1972) when interpreting his findings.

In 1978, Triandis was asked to review Hofstede's (1980) manuscript. Based on a factor analysis of the sum of all of the responses in each culture, Hofstede named one of the factors individualism versus collectivism, and defined it as follows (1980: 51):

Individualism pertains to societies in which the ties between individuals are loose; everyone is expected to look after himself or herself and his or her immediate family. Collectivism as its opposite pertains to societies in which people from birth onwards are integrated into strong, cohesive ingroups, which throughout people's lifetime continue to protect them in exchange for unquestioning loyalty.

Triandis enthusiastically endorsed the book and began a research program which examined the question: How should the collectivism–individualism contrast be defined? In the next decade, a number of studies coming from the University of Illinois sought to elucidate the meaning of the constructs. In the first study, Hui and Triandis (1986) asked a sample of anthropologists and psychologists to provide their insight into the meaning of the concepts. They emphasized the centrality of groups *versus* the centrality of individuals. The next set of studies sought to address the question: How should the constructs be operationalized? The first measurement was provided in Hui's (1988) dissertation which was based on the themes that were identified in Hui and Triandis

(1986), and which eventually became the INDCOL measurement of the construct (for later measures developed in this research program, see Triandis and Gelfand, 1998; Triandis et al., 1986, 1988, 1990, 1995, 1998).

Triandis and his students found general consensus, and based on their results, they developed more items to further investigate the constructs. In a series of studies, Triandis et al. (1986) examined the structure of these items at the culture level in nine countries. Their culture-level analysis revealed four factors, two of which were reflective of individualism (*self-reliance with hedonism* and *separation from ingroups*) and two of which were reflective of collectivism (*family integrity* and *interdependence with sociability*). They found Hofstede's (1980) nation scores on individualism and collectivism were only correlated with scores on family integrity (collectivism) ($r = 0.78$). Triandis et al. (1993) extracted multiple universal (i.e., *etic*) and culture-specific (i.e., *emic*) independent dimensions of individualism and collectivism across cultures. Thus, unlike previous analyses, Triandis and colleagues found evidence of the multidimensionality of the constructs at the culture level, while at the same time confirming the overlap of some of the dimensions with Hofstede's (1980) original work.

At the same time that Triandis was validating new measures of individualism and collectivism, Shalom Schwartz and his [p. 505 ↓] colleagues (Schwartz, 1992; Schwartz and Bilsky, 1990) were exploring universals in individualism and collectivism values. In a study of the structure of values among over 44,000 teachers and students in 54 countries, Schwartz examined the extent to which people view themselves as autonomous versus embedded in groups, reflecting what they referred to as emphasis on *autonomy versus conservation*.

While work by Triandis and Hofstede was largely "bottom up" (emerging from the data), Schwartz predicted *a priori* the nature of value dimensions, such as autonomy (individualism) and conservation (collectivism), as well as the relations among these values and other values in the circumplex (for other large scale studies of individualism–collectivism see the Chinese Culture Connection, 1987; House et al., 2004; and Smith et al., 1996; for more recent measures, see Fischer et al., 2009; Shteynberg et al., 2009; Zou et al., 2009).

Individualism–Collectivism Meets Social Cognition Research: The 1990s and beyond

In the late 1980s and early 1990s, two papers had a profound influence on the direction of individualism–collectivism research by connecting theory and empirical work on the constructs with basic social cognition research on the self. In his *Psychological Review* paper “The self and social behavior in differing cultural contexts,” Triandis (1989) set forth a theory of how different aspects of the self (private, public, collective) are sampled with different probabilities in different kinds of social environments. It was in this paper that Triandis argued that the more individualistic the culture, the more frequent the sampling of the private self and the less frequent the sampling of the collective self. By contrast, he argued that collectivism, external threat, competition with outgroups, and common fate increase the sampling of the collective self. At much the same time, a seminal paper by Markus and Kitayama (1991) also advanced key propositions linking the constructs (which they labeled independence and interdependence) to fundamental cultural differences in cognition, emotion, and motivation. The importance of these papers cannot be underestimated, as they, for the first time, situated the study of individualism and collectivism in the mainstream of social psychology. “Basic” findings on the self that were thought to be universal, whether related to self-efficacy, self-enhancement, self-verification, self-actualization, self consciousness, self-control, among others, were challenged and illustrated to be reflective of Western norms and assumptions of individualism (see Markus and Kitayama, 1991).

The influence of the social cognition movement on research on individualism and collectivism can also be seen in other scholars' works. In a landmark study, Trafimow et al. (1991) showed that, as with other constructs that can be primed, asking people to think for a few minutes if they are similar or different from family and friends increases the probability of collectivist or individualist responses. Likewise, Hong et al. (2000) showed that bicultural individuals could be made to have collectivist or individualist mindsets depending on the primes they have received. In a seminal meta-analysis, Oyserman and Wing-sing Lee (2008) illustrated that individualism and collectivism

Primes have reliable and consistent effects on values, relationality, self-concept, and cognition, across different types of primes and samples. More generally, they showed that the “cognitive tools” that are brought online when collectivism is primed focus on connecting, integrating, and assimilating the figure with the ground and the self with other. By contrast, the “cognitive tools” that are brought online when individualism is primed focus on pulling apart and separating, and contrasting the figure and ground, self and other. Much work by Nisbett and his collaborators (2001) also connected individualism and collectivism research to basic thought processes as discussed below.

[p. 506 ↓]

Defining Attributes and Correlates of Individualism and Collectivism

We began this chapter with the roots that individualism and collectivism had in the *Analysis of Subjective Culture*. In the decades since this book was published, and through many scholarly efforts, much research has documented key defining attributes of the constructs and its antecedents and consequences, which are reviewed below.

Defining Attributes

Research has illuminated a number of defining attributes of individualism and collectivism (Triandis, 1995):

- 1. In collectivistic cultures, the self is interdependent with some group versus in individualistic cultures, the self is independent of groups (Markus and Kitayama, 1991). For example, when asked to complete 20 sentences that begin with “I am...,” in collectivist samples about 35 percent of the responses have “social” content (I am an uncle, I am a member of my fraternity). By contrast, only about 15 percent of the responses from individualist samples have social content. In fact, the mode of social content of 500 Illinois students was zero (Triandis et al., 1990).

- 2. In collectivist cultures, the goals of the group have priority over individual goals and ingroup and individual goals are usually the same; in individualistic cultures, the goals may be different, and if they are in conflict the individual's goals have priority over the goals of the group (Triandis, 1995).
- 3. In collectivist cultures, norms, obligations, and duties guide behavior, whereas in individualist cultures attitudes, personal needs, individual rights, and the contracts the individual has established with others are important determinants of behavior (e.g., Davidson et al., 1976).
- 4. In collectivist cultures, communal relationships (Mills and Clark, 1982) are most frequent, and individuals stay in unpleasant groups or relationships; in individualist cultures individuals tend to leave unsatisfactory relationships (Kim, 1994).

Antecedents of Individualism and Collectivism

Numerous ecological, institutional, situational, and demographic antecedents of individualism and collectivism have been advanced in the literature. As a general principle, factors that increase the need for people to rely on others and which activate common fate promote collectivism. Factors that allow individuals to separate from others promote individualism.

Ecology

Numerous cross-cultural scholars have posited that cultural differences in individualism and collectivism develop as adaptations to the ecological context (Berry, 1976). Collectivism is generally found in agricultural societies wherein conformity and obedience are crucial for survival, whereas more individualism is found among hunters and in complex (e.g., information) societies than in nomadic or agricultural societies wherein self-reliance and freedom are crucial for survival (Barry et al., 1959). An open frontier (Kitayama et al., 2006) also increases the probability of individualism given that it allows people to separate and live at a distance from other people (Triandis, 1995).

Likewise, rural contexts in which there is low mobility and people need to fit into their communities are more collectivistic than urban contexts (Realo et al., 1997). In sum, when the ecology requires connection versus separation, this increases the probability of collectivism versus individualism, respectively.

Family Structure

Family structures that promote embeddedness among individuals promote collectivism whereas family structures that allow separation among individuals promote individualism. Individualism is often associated with nuclear family structures, whereas collectivism is associated with extended family structures (Triandis, 1989). In a 16-culture study, Georgas et al. (2001) found that members of [p. 507 ↓] individualistic cultures lived farther away from grandparents, aunts/uncles, and cousins and visited them less than members of collectivist cultures. In families with many children and therefore greater interdependence, there is a higher probability of collectivism, whereas in families with only children there is a greater probability of individualism (Falbo, 1992). Individualism at the country level is also significantly related to divorce rates (Lester, 1995).

Distribution of Wealth

Wealth affords separation from others and has been associated with higher individualism. Hofstede (1980) found a positive correlation between individualism and wealth, with industrialized wealthy countries scoring higher on individualism than developing countries. Hofstede (1980) later addressed the issue of causality, and argued that an increase in national wealth causes an increase in individualism in a culture, and not vice versa. In this view, individualism is thought to increase as the discretionary capital that is available to people. As people become more affluent, they have more freedom to do their own thing, and accordingly “financial independence leads to social independence” (Triandis, 1994: 165).

Situational Conditions

Situations in which common fate and the need for interdependence are made salient (Campbell, 1958) and in which there are crises and threats to the ingroup (McKelvey, 1982) increase the probability of collectivism. The more people are rewarded for group action, the more likely it is that the culture will be collectivist, whereas the more they are rewarded for individual actions the more likely it is that the culture will be individualist (Lillard, 1998). Disjunctive tasks (that can be accomplished by just one member of a group) increase individualism, whereas conjunctive tasks (that require all members of the group to contribute) increase collectivism (Breer and Locke, 1965).

As noted above, subtle situational priming also affects individualism and collectivism. When the collective self of collectivist participants is primed (by asking people to think of what they have in common with their family and friends, or by exposing participants to words like “we” or “us;” Oyserman and Wing-sing Lee, 2008) participants emit collectivist responses. Collectivist languages, like Chinese or Nepali, can also be used as primes. Samples that are exposed to themes of independence and autonomy in the media are more individualist. For example, exposure to Hollywood-type media increases individualism (McBride, 1998).

Demographics

In general, the lower the status of a group in a social hierarchy, the more likely it is to be collectivistic. Low status requires sharing of resources and the development of values that emphasize security, reliability, and tradition (e.g., Kohn, 1969; Triandis, 2009a). Indeed, research has shown that across many societies, the lower social classes are more collectivist than the upper classes (Kohn, 1969). Schwartz and Smith (1997) also reported that younger and more educated individuals tend to be more individualistic than older and less educated individuals across many societies. In the US, persons of color have scored higher on collectivism (defined as an orientation toward the welfare of one's larger community) and familism (defined as an orientation toward the welfare of one's immediate family) as compared to Caucasians (Gaines et al., 1997; but see

Jones, 1997 for a contrasting analysis). With respect to gender, Kashima et al. (1995) found no difference between males and females across five countries on individualism and collectivism. Gender differences, however, were found for a separate construct: relationality (see also Gabriel and Gardner, 1999).

Consequences of Individualism and Collectivism

Individualism and collectivism have been shown to have wide-ranging consequences [p. 508 ↓] for social-psychological phenomena. As a general principle, collectivism promotes cognitions, motivations, emotions, and behaviors all in the service of connecting with one's group, while individualism promotes cognitions, motivations, emotions, and behaviors all in service of pulling apart and separating from others. Below is a summary of a number of important implications of individualism and collectivism (see Gelfand et al., 2004; Kitayama and Cohen, 2007; Triandis, 1995).²

Focus of Attention

In collectivistic cultures, relationships are the figure, and individuals are in the background; in individualistic cultures, individuals are the figure and groups are in the background.

Attributions

In collectivistic cultures, individuals tend to make external attributions (e.g., norms, roles, group pressure) concerning the determinants of behavior whereas in individualistic cultures, individuals tend to make internal attributions (e.g., attitudes, personality) (Morris and Peng, 1994). In collectivistic cultures, individuals tend to attribute success to the help received from others and failure to a lack of effort. By

contrast, in individualistic cultures, individuals tend to attribute success to their own ability and failure to luck or task difficulty.

Self Definition

In collectivistic cultures, individuals define the self in context and see the environment as more or less fixed and themselves as changeable, whereas individualists see themselves as more or less stable (invariant attitudes, personality, rights) and the environment as changeable (Chiu et al., 1997; Chiu and Hong, 1999). People in collectivistic cultures generally know more about others than about themselves, whereas people in individualistic cultures generally know more about themselves than about others (Markus and Kitayama, 1991). For collectivists, the self includes the achievements of the group; for individualists, the self includes the achievements of the individual.

Goals

Individuals in collectivistic cultures are motivated by others' choices, whereas individuals in individualistic cultures are motivated when they have a personal choice (Iyengar and Lepper, 1999). Likewise, there is more self-efficacy experienced when working alone in individualistic cultures whereas there is more self-efficacy experienced when working in groups in collectivistic cultures (Earley, 1993). People in collectivistic cultures are also more prevention-focused whereas people in individualistic cultures are more promotion focused (Lalwani et al., 2009; Lee et al., 2000).

Emotions

People in collectivistic cultures tend to have more engaged emotions, such as sympathy, while people in individualistic cultures tend to have more disengaged emotions, such as pride (Kitayama et al., 2007). Collectivists express both good and bad feelings, and tend to be moderate in their expression (Mesquita and Leu, 2007).

Cognitions

The more collectivist the culture the more are individuals likely to use holistic and circular thinking and pay attention to the field, whereas the more individualist the culture the more people are likely to use linear and analytic thinking and pay attention to the object (Nisbett, 2003). People in individualist cultures tend to make more judgments based on explicit rules, while people in collectivist cultures tend to make more judgments based on the family resemblance of the various stimuli (Kitayama et al., 2007). Collectivists also tend to use dialectical thinking and are tolerant of contradiction (Peng and Nisbett, 1999).

Norms

Equality and need are emphasized in the distribution of resources among collectivists, particularly with ingroup members, and equity is emphasized among individualists (Leung, 1997). Norms for behavior are more cooperative among collectivists and more [p. 509 ↓] competitive among individualists (Gelfand and Realo, 1999).

Values

The values of collectivists tend to emphasize family security, social order, respect for tradition, harmony, politeness (Schwartz, 1994). Loyalty to the employer and the country are also important (Engel, 1988). The values of individualists tend to emphasize being curious, broadminded, creative, and having a varied and exciting life (Schwartz, 1994). Independence and self-sufficiency are also important (Engel, 1988).

Calamities

In collectivistic cultures, a major calamity is ostracism; in individualistic cultures, a major calamity is dependence on others (Triandis, 1995).

Ingroups

Collectivists have few ingroups and relationships within them are intense. Individualists have many ingroups, and relations are superficial. In collectivistic cultures, self-sacrifice for ingroup is expected and there is cooperation within ingroups; in individualistic cultures, less willingness for self-sacrifice is expected and debate and confrontation are acceptable in ingroups. In collectivistic cultures, the ingroup is perceived as more homogeneous than outgroups, whereas in individualistic cultures, the ingroup is perceived as more heterogeneous than outgroups. In collectivistic cultures, ingroups are defined by similarity to kinship, tribe, religion, race, language, and village whereas in individualistic cultures, ingroups are defined by similarity in achieved attributes (e.g., profession). When making judgments about the trustworthiness of others, people in collectivistic contexts rely on situational signs (e.g., benevolent interactions with the other) whereas people in individualistic cultures tend to rely on dispositional signs (e.g., ability and integrity) (Branzei et al., 2007).

Social Behavior

In collectivistic cultures, behavior is mostly a function of norms and there is a large difference when behavior is toward an ingroup versus an outgroup member. In individualistic cultures, behavior is mostly a function of attitudes and there is less of a distinction between ingroups and outgroups. In the former, people have few skills to enter new groups; in the latter people are skilled in entering and leaving groups. In the former there are communal exchanges and much intimacy. In the latter there are contractual exchanges and less intimacy.

Perceived Determinants of Social Behavior

In collectivistic cultures, ingroup norms, group memberships, context, age, gender, and social relations are especially important determinants of social behavior. In

individualistic cultures, beliefs, attitudes, values, and achieved roles are especially important determinants of social behavior.

Language and Communication

Language and communication in individualistic cultures is direct and emphasizes the individual whereas it is more indirect and de-emphasizes the individual in collectivistic cultures. For example, pronouns such as “I” and “you” are widely used in individualistic cultures and are frequently dropped in collectivistic cultures (Kashima and Kashima, 1998). In individualistic cultures people use more adjectives which suggests more of a dispositional perspective wherein there is low contextual focus. In collectivist cultures they use more action verbs which suggest more of a contextual and situated focus (Zwier, 1998). The more collectivist the culture the more people are likely to communicate indirectly (paying attention to gestures, body position, tone of voice, and loudness of voice) (Holtgraves, 1997; Triandis, 1994).

Group Processes

Individualism is associated with general resistance to teams at the individual and [p. 510 ↓] group level of analysis (Kirkman and Shapiro 2001a, 2001b). Individuals in collectivistic cultures are more likely to perceive groups as “entities” which have agentic qualities and dispositions as compared with individuals in individualistic cultures (e.g., Chiu et al., 2000, Kashima et al., 2005, Morris et al., 2001). Collectivism is associated with greater conformity (Bond and Smith, 1996), cooperation (Cox et al., 1991; Eby and Dobbins, 1997; Wagner, 1995), and more organizational citizenship behaviors (i.e., prosocial behaviors) (Moorman and Blakely, 1995). Schemas for what constitutes “successful” workgroups also vary across cultures. Collectivists perceive that socioemotional behaviors are important for group success, whereas individualists perceive that high task orientation and low socioemotional behaviors are important for group success (Sanchez-Burks et al., 2000). Individuals also hold groups and organizations more accountable for failed actions in collectivistic cultures as compared

to individualistic cultures (Chiu and Hong, 1992; Chiu et al., 2000; Menon et al., 1999; Zemba et al., 2006).

Leadership

The attributes that are perceived as important for effective leadership vary across individualistic and collectivistic societies (House et al., 2004). Collectivistic societies and organizations are more likely to endorse charismatic leadership, self-protective leadership (i.e., face-saving leader behaviors), and team-oriented leadership as compared with individualistic societies (Gelfand et al., 2004). Also, leaders' behavior is interpreted differently depending on the culture. For example, "talking behind one's subordinates back" is perceived to be negatively related to considerate leadership in the US, where it is seen as inappropriate to indirectly speak to one's employees. However, such behaviors are positively related to consideration in Japan where face saving and indirect communication are seen as important (Smith et al., 1989).

Conflict and Negotiation

Individualism and collectivism affect how individuals perceive and manage conflict. Individualists perceive conflicts to be more about violations of individual rights and autonomy whereas collectivists perceive the same conflicts to be about violations of duties and obligations (Gelfand et al., 2001). Negotiators in individualistic cultures tend to be susceptible to host of competitive biases in negotiations, including self-serving biases (Gelfand et al., 2002), fixed-pie biases (Gelfand and Christakopoulou, 1999), and dispositional attribution biases as compared to negotiators in collectivistic cultures (Morris et al., 1999). Negotiators tend to share information directly (e.g., through questions about preferences) in individualistic cultures whereas they tend to share information indirectly (through offer behavior) in collectivistic cultures (e.g., Adair et al., 2001). Situational factors, such as being accountable to constituents (Gelfand and Realo, 1999) or having a high need for closure (Fu et al., 2007, Morris and Fu, 2001) amplify cultural differences in conflict and negotiation.

Additional Distinctions: Vertical and Horizontal Dimensions of Individualism and Collectivism

Triandis (1995) argued that there are many “species” of individualism and collectivism. One such distinction – which was found also in the original *Analysis of Subjective Culture* – is that individualism and collectivism can be both horizontal and vertical. In vertical cultures individuals are motivated to stand out. In horizontal cultures individuals avoid standing out (they try to blend in) (Daun, 1992). Traditional India is vertical, while Australia and Sweden are horizontal. In India individuals seek status, and figuratively they want to stand out, to be “on top of an elephant” parading the streets to the applause of an adoring population. On the other hand, in Australia tall poppies are brought down (Feather, 1994) and in Sweden people avoid standing out (Daun, 1992).

[p. 511 ↓]

The major value of vertical individualists is achievement. Americans are offended if someone tells them that they are “average” (Weldon, 1984) which suggests considerable vertical individualism. Vertical individualism increases the probability of competition (Triandis, 1995). In vertical individualist cultures people are high in the need for power, achievement, and prestige (Daun, 1992). The major value of horizontal individualists is uniqueness (Triandis and Gelfand, 1998). Horizontal individualism increases the probability that individuals will be motivated by (a) the good, comfortable life, and (b) will seek to be unique without standing out (Inglehart, 1997). The major value of horizontal collectivists is cooperation. The Israeli kibbutz is an example of such a culture. In horizontal collectivist cultures people are high in the need for affiliation, and in modesty (Kurman, 2001, 2003). The major concern of vertical collectivists is to do their duty. Traditional cultures (e.g., Indian village) are high in this tendency. In vertical collectivist cultures people are motivated, more than in other cultures, to conform to authorities (Bond and Smith, 1996). While there are considerable cross-cultural differences on the constructs, it is also important to point out that individuals have all four of the cognitions (Triandis et al., 1998). There is now a literature (e.g.,

Choiu, 2001; Kurman and Sriram, 2002; Nelson and Shavitt, 2002; Soh and Leong, 2002) that shows that vertical collectivists are quite different from horizontal collectivists and vertical individualists are different from horizontal individualists.

Cultures may emphasize, at particular time periods, a particular syndrome more than the other syndromes. For example, Galtung (1979) divided the history of Europe into three parts: Antiquity up to the fall of the Roman Empire (476 AD), which was dominated by vertical individualism; the Middle Ages, up to the fall of Constantinople in 1453, which was characterized by vertical collectivism; and the modern period, which was characterized by vertical individualism, with the exception of Scandinavian cultures which tend to be higher on horizontal individualism (Triandis, 1995).

Evaluation of Individualism and Collectivism

Both individualism and collectivism have positive and negative effects on societies and individuals therein. There are data suggesting that divorce, delinquency, drug abuse, heart attacks, and suicide are higher in individualist than in collectivist cultures (Eckersley and Dear, 2002; Triandis et al., 1988). Eckersley, an Australian epidemiologist, argues that individualism is undesirable from the point of view of mental health. Torrey and Miller (2001) reported that the number of insane persons per 1,000 has increased steadily since the beginning of the industrial revolution in England, Ireland, Canada, and the US. The four curves that cover the 1807–1961 period (when insane people were placed in communities in all four countries, so that there is no longer any reliable measurement of this rate) are impressively steep. During this period there have been increases in both affluence (i.e., cultural complexity), and looseness. Thus, theoretically, there has been an increase in individualism. The authors hypothesize that living in cities, changes in diet, alcohol consumption, more toxins in the environment, improved medical care that does not eliminate unfit babies, infectious agents or a combination of these factors might account for the fact that the rates increased seven-fold between 1750 and the present. While there are likely different definitions of mental illness across the world, it is important to consider this work in evaluating the constructs.

On the other hand, subjective wellbeing is higher in individualist than in collectivist cultures (Diener et al., 1995). The US is thirteenth in the world on subjective wellbeing (Tov and Diener, 2007). Japan, which is [p. 512 ↓] generally tighter and more collectivist than the United States, scored thirty-fifth. Individualism is also related to high life expectancy, higher satisfaction, and higher scores on the Human Development index (see Gelfand et al., 2004).

The picture is complex and difficult to evaluate. Individualism is a desirable cultural pattern for those who want to achieve, to become distinguished; collectivism is a desirable cultural pattern for those who want to be embedded in social relationships. Longevity is higher in Japan (82) than in most of the West. "Life without disease" averages 74.5 years in Japan, 73.2 in Australia, 73.1 in France, 72.8 in Spain, 72.7 in Italy, 72.5 in Greece, 72.5 in Switzerland, and is only 70.0 in the US (*The Economist*, 2009). It maybe that extreme individualism is not desirable, and a culture that has both collectivist and individualist elements is ideal.

Practical Implications of Individualism and Collectivism

As a major dimension of cultural variation, individualism and collectivism has much practical relevance for managing interdependence in an increasingly "flat" world. The sheer amount of intercultural contact across many areas of life is unprecedented. For example, it is estimated that over 175 million people migrated across national borders each year¹ and 1 in every 35 people in the world lives in a country different than their birth (Office of the United Nations High Commissioner for Human Rights, 2009). Globalization has dramatically increased contact among people of different cultures, with new mergers, acquisitions, and global ventures happening on a daily basis and millions of expatriates crossing boarders every year. Pressing global concerns, such as preserving the environment, require intercultural collaboration at the international and local levels. And in a world of global threats of conflict and terrorism, understanding cultural differences can arguably be a matter of life or death. Indeed, in the recent Iraqi Study report, James Baker noted that understanding cultural differences on behavior is of the highest national priority (Baker and Hamilton, 2006).

Within this global context, it is hard to underestimate the importance of the knowledge that has been gleaned from last four decades of individualism and collectivism research. Cross-cultural differences along the individualism and collectivism divide, if not understood and managed, can have disastrous consequences across many domains of life. Within organizational contexts, a lack of understanding of cultural differences can translate into failed mergers and ventures, premature expatriate return and turnover, and organizational conflict. In the US alone, for example, more than 100,000 companies conduct business overseas, and at least one-third of American profits are derived from international business dealings (Erez, 1994); thus, understanding key cultural differences related to individualism and collectivism is a key business imperative.

The need to understand individualism and collectivism to help negotiators negotiate effectively across cultures is also painfully obvious in today's geopolitical scene, where the source of conflict among humankind is thought to be increasingly cultural in nature (Huntington, 1996). Anecdotal examples abound of failed intercultural negotiations, many of which can be linked to basic differences in individualism and collectivism (Gelfand and Realo, 1999; Triandis, 1994). For example, during a summit in 1969, Prime Minister Sato of Japan was told by President Nixon for Japan to exercise export restraint to which he responded, "zenso shimasu" ("I will do my best"). Although Sato really meant "no," Nixon misunderstood this to mean agreement, and when there was no implementation, Nixon denounced Sato as a liar (Cohen, 1997). Without understanding of core elements of individualism and collectivism, such as indirectness versus directness in communication, as reviewed above, intercultural negotiations are seriously compromised. [p. 513 ↓] As another practical example, in the area of health, a failure to understand individualism and collectivism can put patients at serious risk. Much research has relevance of cultural factors to health behaviors such as symptom recognition and help-seeking (e.g., Fabrega, 1994; Kleinman et al., 1978; Zola, 1966). In this respect, doctors' understanding of core cultural differences is critical in the treatment of disease.

With Lewin's famous adage that there is nothing as practical as a good theory in mind, cross-cultural psychologists have been translating the knowledge gained over the last four decades on individualism and collectivism into training programs. As early as the *Analysis of Subjective Culture*, knowledge of individualism and collectivism was developed into "cultural assimilator" training programs, which are books, or computer-

based cross-cultural training devices (Fiedler et al., 1971). Assimilators consist of about 100 episodes that reflect a problematic interaction between members of two cultures. For example, an episode might be that an American teacher notices that a Hispanic child does not look at her when she is talking. Under each episode, in the format of a multiple-choice test, are four or five attributions that could explain what is happening in the episode. The trainee is asked to select one of the attributions, and then receives feedback. Some of the attributions are incorrect, and when the trainee selects one of them he/she is asked to try again. When the correct attribution is selected the feedback explains the cultural difference. In the example above the cultural difference is that in the US children are expected to look at a teacher who is talking; but in Latino countries one is “insolent” if one looks directly at a high status person in the eye, and the proper behavior is to respectfully look down.

The attributions are pre-tested with samples from the two cultures, and when members of the host culture select an attribution that is rarely selected by members of the trainee's culture, they begin to understand why the response is not “culturally” accurate. The effect of this training is that trainees learn to make “*isomorphic attributions*,” that is, attributions that are more or less like the attributions that are usually made by members of the host culture in the particular situation. Trainees randomly assigned to a training and a no-training condition show some improvement. The trained feel more comfortable when they are in the host culture.

Early assimilators were developed with a pair of cultures in mind, generally with the aim of training Americans to live in other cultures (e.g., Thailand, Honduras, Japan, Venezuela). For example, the Japanese culture assimilator has 57 incidents that are divided into themes (e.g., hierarchy, face saving behaviors, harmony and emotional control, group-related behaviors, and norm-related behaviors). Typical learnings include such do's and don'ts such as “Students do not wear jewelry to school,” “Demeaning oneself is a proper behavior,” “Newcomers give small gifts,” and “Do not criticize supervisors,” among others (Bhawuk, 2001). More recently, however, it has been found that people learn better when the episodes provide feedback organized around cultural syndromes such as individualism and collectivism. By organizing assimilators around elements of individualism–collectivism, trainees are able to move beyond “do's and don'ts” in other cultures – typically superficial attributes – to understand cultural differences through a coherent framework. Bhawuk (2001) outlined such an approach,

which includes critical incidents that capture the four defining features of individualism and collectivism reviewed above (Triandis, 1995), including the nature of the self, goal priorities, predictors of behavior, and nature of relationships, along with critical incidents that tap into horizontal and vertical dimensions of the constructs. Importantly, however, although people learn much about another culture through cultural assimilators, they do not change their behavior enough to be really successful in the other culture. Changes in behavior require clinical interventions and [p. 514 ↓] behavior shaping. Nonetheless, this application is reported in several handbooks concerned with cross-cultural training (Landis and Bhagat, 1996).

Conclusion

Triandis (2009b) discussed the factors used by individuals to construct the way they see the world. Culture is one of them. Much international conflict is due to differences in the subjective cultures of various groups. The future of the planet may depend on further analyses of the subjective cultures around the globe in order to increase understanding, promote wellbeing, and manage our increasing interdependence.

Acknowledgments

This research is based upon work supported by the US Army Research Laboratory and the US Army Research Office under grant number W911NF-08-1-0144.

Notes

1 The terms *emic* and *etic* in cross-cultural psychology have been derived from linguistics, where *emic* refers to sounds specific to a particular language and *etic* refers to sounds that are universal to languages.

2 Important caveats, however, are in order. First, a distinction needs to be made between data that use “culture” as the unit of analysis (e.g., Hofstede, 1980), and data that use “individuals” as the unit of analysis. When culture is the unit of analysis the

sum or the responses of individuals are entered and the correlation is across cultures. When individuals are the unit of analysis the correlation is within one culture across individuals. Collectivism tends to be the opposite of individualism when culture is the unit of analysis. However, when individuals are the units of analysis, the tendencies toward individualism and collectivism can be orthogonal to each other (Gelfand et al., 1996). A person can be high in both attributes, or high in one and low on the other attribute. The best way to conceptualize this is to think that both collectivist and individualist cognitions are present in every individual, and the tendencies are elicited by the situation.

Harry C.TriandisMichele J.Gelfand

References

- Adair, W.L., Okimura, T. and Brett, J.M. Negotiation behavior when cultures collide: The United States and Japan . Journal of Applied Psychology , 86, 371–385. (2001) <http://dx.doi.org/10.1037/0021-9010.86.3.371>
- Bakan, D. (1996) The Duality of Human Existence . Chicago: Rand McNally.
- Baker, J.A. and Hamilton, L.H. (2006) The Iraq Study Group Report: The Way Forward — A New Approach . US Institute of Peace . New York: Vintage Books.
- Barry, H., III, Child, I.L. and Bacon, M.K. A cross-cultural survey of sex differences in socialization . Journal of Abnormal and Social Psychology , 55, 327–332. (1959)
- Bellah, R.N., Madsen, E., Sullivan, W., Swidler, A. and Tipton, S.M. (1985) Habits of the Heart: Individualism and Commitment in American Life . Berkeley: University of California Press.
- Berry, J.W. Sex differences in behavior and cultural complexity . Indian Journal of Psychology , 51, 89–97. (1976)

Bhawuk, D.P.S. Evolution of cultural assimilators: Toward theory-based assimilators . International Journal of Intercultural Relations , 25, 141–163. (2001) <http://dx.doi.org/10.1016/S0147-1767%2800%2900048-1>

Bond, R. and Smith, P.B. Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task . Psychological Bulletin , 119, 111–137. (1996) <http://dx.doi.org/10.1037/0033-2909.119.1.111>

Branzei, O., Vertinsky, I. and Camp, R.D. Culture contingent signs of trust in emerging relationships . Organizational Behavior and Human Decision Processes , 104, 61–82. (2007) <http://dx.doi.org/10.1016/j.obhdp.2006.11.002>

Breer, P.E. and Locke, E. (1965) Task Experience as a Source of Attitudes . Homewood, IL: Dorsey Press.

Brewer, M.B. and Chen, Y. Where (and who) are collectives in collectivism: Toward conceptual clarification on individualism and collectivism . Psychological Review , 114, 133–151. (2007) <http://dx.doi.org/10.1037/0033-295X.114.1.133>

Brewer, M.B. and Yuki, M. (2007) Culture and social identity . In S. Kitayama, ed. and D. Cohen (eds), Handbook of Cultural Psychology , pp. 307–322. New York: Guilford Press.

Campbell, D.T. Common fate, similarity, and other indices of status of aggregates of persons as social entities . Behavioral Science , 3, 14–25. (1958) <http://dx.doi.org/10.1002/bs.3830030103>

Carneiro, R.L. (1970) Scale analysis, evolutionary sequences, and the ratings of cultures . In R. Naroll, ed. and R. Cohen (eds), A Handbook of Method in Cultural Anthropology . New York: Columbia University Press.

Carpenter, S. Effects of cultural tightness and collectivism on self-concept and causal attribution . Cross-Cultural Research , 34, 38–56. (2000) <http://dx.doi.org/10.1177/106939710003400103>

Chen, F.F. and West, S.G. Measuring individualism and collectivism: The importance of considering differential components, reference groups, and measurement invariance . Journal for Research on Personality , 42, 259–294. (2008) <http://dx.doi.org/10.1016/j.jrp.2007.05.006>

Chiao, J. and Ambady, N. (2007) Cultural neuroscience: Parsing universality and diversity across levels of analysis . In S. Kitayama, ed. and D. Cohen (eds), The Handbook of Cultural Psychology , pp. 237–254. New York: Guilford Press.

Chick, G. Cultural complexity: The concept and its measurement . Cross-Cultural Research , 31, 275–307. (1997) <http://dx.doi.org/10.1177/106939719703100401>

Chinese Culture Connection Chinese values and the search for culture-free dimensions of culture . Journal of Cross-Cultural Psychology , 18, 143–164. (1987) <http://dx.doi.org/10.1177/0022002187018002002>

Chircov, V.I., Lynch, M. and Niwa, S. Application of scenario questionnaire of horizontal and vertical individualism and collectivism to the assessment of cultural distance and cultural fit . International Journal of Intercultural Relations , 29, 469–490. (2005) <http://dx.doi.org/10.1016/j.ijintrel.2005.05.014>

Chiu, C., Dweck, C.S., Tong, J.Y. and Fu, J.H. Implicit theories and concepts of morality . Journal of Personality and Social Psychology , 73, 923–940. (1997) <http://dx.doi.org/10.1037/0022-3514.73.5.923>

Chiu, C. and Hong, Y. The effect of intentionality and validation of collective responsibility attribution among Hong Kong Chinese . Journal of Psychology , 126, 291–300. (1992)

Chiu, C. and Hong, Y. Social identification in a political transition: The role of implicit beliefs . International Journal of Intercultural Relations , 23, 297–218. (1999) <http://dx.doi.org/10.1016/S0147-1767%2898%2900040-6>

Chiu, C., Morris, M.W., Hong, Y.Y. and Menon, T. Motivated cultural cognition: The impact of implicit cultural theories on dispositional attribution varies as a function of

need for closure . *Journal of Personal and Social Psychology* , 78, 247–259. (2000) <http://dx.doi.org/10.1037/0022-3514.78.2.247>

Choi, J. Horizontal and vertical individualism and collectivism among college students in the United States, Taiwan, and Argentina . *Journal of Social Psychology* , 141, 667–678. (2001) <http://dx.doi.org/10.1080/00224540109600580>

Cohen, R. (1997) Negotiating Across Cultures: International Communication in an Interdependent World . Washington, DC: United States Institute of Peace Press.

Cole, M. (1996) Cultural Psychology: A Once and Future Discipline . Cambridge, MA: Harvard University Press.

Cox, T.H., Lobel, S.A. and McLeod, P.L. Effects of ethics group cultural differences on cooperative and competitive behavior on a group task . *Academy of Management Journal* , 34, 827–847. (1991) <http://dx.doi.org/10.2307/256391>

Daun, A. (1992) Modern and modest. Mentality and self-stereotypes among Swedes . In A. Sjoegren, ed. and L. Janson (eds), Culture and Management . Stockholm: Institute of International Business, Series A-7.

Davidson, A.R., Jaccard, J.J., Triandis, H.C., Morales, M.L. and Diaz-Guerrero, R. Cross-cultural model testing: Toward a solution of the etic emic dilemma . *International Journal of Psychology* , 11, 1–13. (1976) <http://dx.doi.org/10.1080/00207597608247343>

Dawood, N.J. (1956) The Koran . New York: Penguin.

Dewey, J. (1930) Individualism Old and New . New York: Minton, Balch.

Diener, E., Diener, M. and Diener, C. Factors predicting the subjective well-being of nations . *Journal of Personality and Social Psychology* , 69, 851–864. (1995) <http://dx.doi.org/10.1037/0022-3514.69.5.851>

Dumont, L. (1986) Essays on Individualism . Chicago: University of Chicago Press.

Durant, W. (1935) Our Oriental Heritage . New York: Simon & Schuster.

- Durkheim, E. (1933) *The Division of Labor*. Chicago: Free Press.
- Earley, C. East meets West, meets Mideast: Further explorations of collectivistic and individualistic workgroups. *Academy of Management Journal*, 36, 319–348. (1993) <http://dx.doi.org/10.2307/256525>
- Eby, L.T. and Dobbins, G.H. Collectivistic orientation in teams: An individual and group-level analysis. *Journal of Organizational Behavior*, 18, 275–295. (1997) <http://dx.doi.org/10.1002/%28SICI%291099-1379%28199705%2918:3%3C275::AID-JOB796%3E3.0.CO;2-C>
- Eckersley, R. and Dear, K. Cultural correlates of youth suicide. *Social Science and Medicine*, 55, 1893–1906. (2002)
- Engel, J.W. Work values of American and Japanese men. *Journal of Social Behavior and Personality*, 3, 191–200. (1988)
- Erez, M. (1994) Toward a model of cross-cultural industrial and organizational psychology. In H. Triandis, ed., M. Dunnette, ed. and L. Hough (eds), *Handbook of Industrial and Organization Psychology*, Vol. 4, 2nd Edition, pp. 559–607. Palo Alto, CA: Consulting Psychologists Press.
- Fabrega, H. International systems of diagnosis in psychiatry. *Journal of Nervous and Mental Disease*, 182, 256–263. (1994) <http://dx.doi.org/10.1097/00005053-199405000-00002>
- Falbo, T. (1992) Social norms and the one-child family. In F. Boer, ed. and J. Dunn (eds), *Children's Sibling Relationship: Developmental and Clinical Issues*, pp. 71–82. Hillsdale, NJ: Erlbaum.
- Feather, N.T. (1994) Attitudes toward high achievers and reactions to their fall: Theory and research concerning tall poppies. In M. Zanna (ed.), *Advances in Experimental Social Psychology*, 25, 1–73. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960151-3>

Fiedler, F.E., Mitchell, T. and Triandis, H.C. The culture assimilation: An approach to cross-cultural training . *Journal of Applied Psychology* , 55, 95–102. (1971) <http://dx.doi.org/10.1037/h0030704>

Fischer, R., Ferrera, M.C., Redford, P., Harb, C., GlazerS., Chen, B-S. et al. Individualism and collectivism as descriptive norms: Development of subjective norms approach to culture measurement . *Journal of Cross-Cultural Psychology* , 40, 187–213. (2009) <http://dx.doi.org/10.1177/0022022109332738>

Fishbein, M. and Ajzen, I. (1975) Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research . Reading, MA: Addison-Wesley.

Fishbein, M., Triandis, H.C., Kanfer, F.H., Becker, M., Middlestadt, S.E. and Eichler, A. (2001) Factors influencing behavior and behavior change . In A. Baum, ed. , T.A. Revenson, ed. and J.E. Singer (eds), *Handbook of Health Psychology* , pp. 3–18. Mahwah, NJ: Lawrence Erlbaum.

Fu, J.H.-Y., Chiu, C.-Y., Morris, M.W. and Young, M.J. Spontaneous inferences from cultural cues: Varying responses of cultural insiders and outsiders . *Journal of Cross-Cultural Psychology* , 38, 58–75. (2007) <http://dx.doi.org/10.1177/0022022106295443>

Gabriel, S. and Gardner, W.L. Are there ‘his’ and ‘hers’ types of interdependence? The implication of gender differences in collective versus relational interdependence for affect, behavior, and cognition . *Journal of Cross-Cultural Psychology* , 28, 642–655. (1999).

Gaines, S.O., Marellich, W.D., Bledsoe, K.L., Steers, W.N., Henderson, M.C., Granrose, C.S. et al. Links between race/ethnicity and cultural values as mediated by racial/ethnic identity and moderated by gender . *Journal of Personality and Social Psychology* , 72, 1460–1476. (1997). <http://dx.doi.org/10.1037/0022-3514.72.6.1460>

Galtung, J. (1979) On the last 2,500 years of Western history . In P. Burke (ed.), *New Cambridge Modern History, Comparison Volume* . Cambridge: Cambridge University Press.

Gelfand, M.J., Bhawuk, D.P.S., Nishii, L.H. and Bechtold, D.J. (2004) Individualism and collectivism . In R.J. House, ed. , P.J. Hanges, ed. , M. Javidan, ed. , P.W. Dorfman, ed. and V. Gupta (eds), Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies , pp. 437–512. Thousand Oaks, CA: Sage.

GelfandM.J. and Christakopoulou, S. Culture and negotiator cognition: Judgment accuracy and negotiation processes in individualistic and collectivistic cultures . Organizational Behavior and Human Decision Processes , 79, 248–269. (1999) <http://dx.doi.org/10.1006/obhd.1999.2845>

Gelfand, M.J., Higgins, M., Nishii, L., Raver, J., Dominguez, A., Yamaguchi, S. et al. Culture and egocentric biases of fairness in conflict and negotiation . Journal of Applied Psychology , 87, 833–845. (2002) <http://dx.doi.org/10.1037/0021-9010.87.5.833>

Gelfand, M.J., Nishii, L. and Raver, J.L. On the nature and importance of cultural tightness-looseness Journal of Applied Psychology , 91, 1225–1244. (2006) <http://dx.doi.org/10.1037/0021-9010.91.6.1225>

Gelfand, M.J., Nishii, L.H., Holcombe, K., Dyer, N., Ohbuchi, K. and Fukumo, N. Cultural influences on cognitive representation of conflict: Interpretation of conflict episodes in the U.S. and Japan . Journal of Applied Psychology , 86, 1059–1074. (2001) <http://dx.doi.org/10.1037/0021-9010.86.6.1059>

Gelfand, M.J. and Realo, A. Individualism-collectivism and accountability in intergroup negotiations . Journal of Applied Psychology , 84, 721–736. (1999) <http://dx.doi.org/10.1037/0021-9010.84.5.721>

Gelfand, M.J., Triandis, H.C. and Chan, D.K.-S. Individualism versus collectivism or authoritarianism? European Journal of Social Psychology , 26, 397–410. (1996) <http://dx.doi.org/10.1002/%28SICI%291099-0992%28199605%2926:3%3C397::AID-EJSP763%3E3.0.CO;2-J>

Georgas, J., Mylonas, K., Bafoiti, T., Poortinga, Y.A., Christakopoulou, S., Kagitcibas, C. et al. Functional relationships in the nuclear and extended family: A 16-culture study . International Journal of Psychology , 36, 289–300. (2001) <http://dx.doi.org/10.1080/00207590143000045>

Glenn, E. and Glenn, P. (1981) *Man and Mankind: Conflicts and Communication between Cultures* . Norwood, NJ: Ablex.

Harb, C. and Smith, P.B. Self-construals across cultures: Beyond independence and interdependence . *Journal of Cross-Cultural Psychology* , 39, 178–197. (2008) <http://dx.doi.org/10.1177/0022022107313861>

Herskovits, M.J. (1955) *Cultural Anthropology* . New York: Knopf.

Hofstede, G. (1980) *Culture's Consequences* . Thousand Oaks, CA: Sage.

Holtgraves, T. Styles of language use: Individual and cultural variability in conversational indirectness . *Journal of Personality and Social Psychology* 73, 624–637. (1997) <http://dx.doi.org/10.1037/0022-3514.73.3.624>

Hong, Y.Y., Morris, M., Chiu, C.-Y. and Benet-Martinez, V. Multicultural minds: A dynamic constructivist approach to culture and cognition . *American Psychologist* , 55, 709–720. (2000) <http://dx.doi.org/10.1037/0003-066X.55.7.709>

House, R.J., Hanges, P.J., Javidan, M., Dorfman, P.W. and Gupta, V. (2004) *Culture, Leadership, and Organizations: The GLOBE Study of 62 cultures* . Thousand Oaks, CA: Sage.

Hui, C.H. Measurement of individualism-collectivism . *Journal of Research on Personality* , 22, 17–36. (1988) <http://dx.doi.org/10.1016/0092-6566%2888%2990022-0>

Hui, C.H. and Triandis, H.C. Individualism-collectivism: A study of cross-cultural researchers . *Journal of Cross-Cultural Psychology* , 17, 225–248. (1986) <http://dx.doi.org/10.1177/0022002186017002006>

Huntington, S.P. (1996) *The Clash of Civilizations and the Remaking of World Order* . New York: Simon & Schuster.

Inglehart, R. (1997) *Modernization and Post Modernization: Cultural, Economic, and Political Change in 43 Societies* . Princeton, NJ: Princeton University Press.

Inglehart, R. and Baker, W.E. Modernization, cultural change, and the persistence of traditional values . American Sociological Review , 65, 19–51. (2000) <http://dx.doi.org/10.2307/2657288>

Iyengar, S.S. and Lepper, M.R. Rethinking the value of choice: A cultural perspective on intrinsic motivation . Journal of Personality and Social Psychology , 76, 349–366. (1999) <http://dx.doi.org/10.1037/0022-3514.76.3.349>

Jackson, C.L., Colquitt, J.A., Wesson, M.J. and Zapata-Phelan, C.P. Psychological collectivism: a measurement validation and linkage to group member performance . Journal of Applied Psychology , 91, 884–899. (2006) <http://dx.doi.org/10.1037/0021-9010.91.4.884>

Kagan, D. (1966) The Ancient Near East and Greece . London: Macmillan Group.

Kashima, E.S. and Hardie, E.A. The development and validation of the Relational, Individual, and Collective Self Aspects (RIC) scale . Asian Journal of Social Psychology , 3, 19–48. (2000) <http://dx.doi.org/10.1111/1467-839X.00053>

Kashima, E.S. and Kashima, Y. Culture and language: The case of cultural dimensions of personal pronoun use . Journal of Cross-Cultural Psychology , 29, 461–486 (1998) <http://dx.doi.org/10.1177/0022022198293005>

Kashima, Y., Kashima, E.S., Chiu, C.-Y., Farsides, T., Gelfand, M., Hong, Y.Y. et al. Culture, essentialism, and agency: Are individuals universally believed to be more real entities than groups? European Journal of Social Psychology , 35, 147–169. (2005) <http://dx.doi.org/10.1002/ejsp.237>

Kashima, Y., Yamaguchi, S., Kim, U., Choi, S.C., Gelfand, M.J. and Yuki, M. Culture, gender and self: A perspective from individualism-collectivism research . Journal of Personality and Social Psychology , 59, 925–937. (1995) <http://dx.doi.org/10.1037/0022-3514.69.5.925>

Kateb, G. (1992) The Inner Ocean: Individualism and Democratic Culture . Ithaca, NY: Cornell University Press.

Kim, U. (1994) Individualism and collectivism: Conceptual clarification and elaboration . In U. Kim, ed. , H.C. Triandis, ed. , C. Kagitcibasi, ed. , S.C. Choi, ed. and G. Yoon. (eds), Individualism and Collectivism Theory, Method, and Application , pp. 19–40. Thousand Oaks, CA: Sage.

Kirkman, B.L. and Shapiro, D.L. The impact of cultural values on job satisfaction and organizational commitment in self-managing work teams: The mediating role of employee resistance . Academy of Management Journal , 44, 557–569. (2001a) <http://dx.doi.org/10.2307/3069370>

Kirkman, B.L. and Shapiro, D.L. The impact of employee cultural values on productivity, cooperation, and empowerment in self-managing work teams . Journal of Cross-Cultural Psychology , 32, 597–617. (2001b) <http://dx.doi.org/10.1177/0022022101032005005>

Kitayama, S. and Cohen, D. (2007) Handbook of Cultural Psychology . New York: Guilford Press.

Kitayama, S., Ishii, K., Imada, T., Takemura, K. and Ramaswamy, J. Voluntary settlement and the spirit of independence: Evidence from Japan's 'Northern frontier.' Journal of Personality and Social Psychology , 91, 369–384. (2006) <http://dx.doi.org/10.1037/0022-3514.91.3.369>

Kleinman, A., Eisenberg, L. and Good, B. Culture, illness, and care: Clinical lessons from anthropological and cross-cultural research . Annals of Internal Medicine , 88, 251–258. (1978)

Kluckhohn, C. (1954) Culture and behavior . In G. Lindzey (ed.), Handbook of Social Psychology , 2, 921–976. Cambridge, MA: Addison-Wesley.

Kluckhohn, C. (1956) Toward a comparison of value emphasis in different cultures . In L.D. White (ed.), The State of the Social Sciences , pp. 116–132. Chicago: University of Chicago Press.

Kluckhohn, F.R. and Strodtbeck, F. (1961) Variations in Value Orientation . Evanston, IL: Row, Peterson & Co.

Kohn, M.K. (1969) Class and Conformity . Homewood, IL: Dorsey Press.

Kurman, J. Is self-enhancement related to modesty or to individualism-collectivism? A test with four Israeli groups . Asian Journal of Social Psychology , 4, 225–238. (2001) <http://dx.doi.org/10.1111/1467-839X.00087>

Kurman, J. Why is self-enhancement low in certain collectivist cultures? Journal of Cross-Cultural Psychology , 34, 496–510. (2003) <http://dx.doi.org/10.1177/0022022103256474>

Kurman, J. and Sriram, N. Interrelationships between vertical and horizontal collectivism, modesty, and self-enhancement . Journal of Cross-Cultural Psychology , 33, 71–86. (2002). <http://dx.doi.org/10.1177/0022022102033001005>

Lalwani, A.K., Shrum, L.J. and Chiu, C-Y. Motivated response style: The role of cultural values, regulatory focus, and self-consciousness in socially desirable responding . Journal of Personality and Social Psychology , 96, 870–882. (2009) <http://dx.doi.org/10.1037/a0014622>

Landis, D. and Bhagat, R.S. (1996) Handbook of Intercultural Training , 2nd Edition. Thousand Oaks, CA: Sage. <http://dx.doi.org/10.4135/9781452231129>

Lee, A.Y., Aaker, J.L. and GardnerW.L. The pleasures and pains of distinct self-construals: The role of interdependence in regulatory focus . Journal of Personality and Social Psychology , 78, 1122–1134. (2000) <http://dx.doi.org/10.1037/0022-3514.78.6.1122>

Lester, D. Individualism and divorce . Psychological Reports , 76, 258. (1995) <http://dx.doi.org/10.2466/pr0.1995.76.1.258>

Leung, K. (1997) Negotiation and reward allocations across cultures . In P.C. Earley, ed. and M. Erez (eds), New Perspectives on International Industrial and Organizational Psychology , pp. 640–675. San Francisco, CA: Lexington Press.

Lillard, A. Ethnopsychologies: Cultural variations in theories of mind . Psychological Bulletin , 123, 3–32. (1998) <http://dx.doi.org/10.1037/0033-2909.123.1.3>

Lomax, A. and Berkowitz, N. The evolutionary taxonomy of culture . *Science* , 177, 228–239. (1972) <http://dx.doi.org/10.1126/science.177.4045.228>

Markus, H. and Kitayama, S. Culture and self: Implications for cognition, emotion, and motivation . *Psychological Review* , 98, 225–253. (1991) <http://dx.doi.org/10.1037/0033-295X.98.2.224>

McBride, A. Television, individualism, and social capital . *Political Science and Politics* , 31, 542–555. (1998) <http://dx.doi.org/10.2307/420614>

McKelvey, B. (1982) *Organizational Systematics: Taxonomy Classification and Evolution* . Los Angeles, CA: University of California Press.

Mead, M. (1967) *Cooperation and Competition Among Primitive Peoples* . Boston: Beacon Press.

Menon, T., Morris, M.W., Chiu, C.-Y. and Hong, Y.Y. Culture and the construal of agency: Attribution to individual versus group disposition . *Journal of Personal and Social Psychology* , 76, 701–717. (1999) <http://dx.doi.org/10.1037/0022-3514.76.5.701>

Mesquita, B. and Leu, J. (2007) Cultural psychology of emotion . In S. Kitayama, ed. and D. Cohen (eds), *Handbook of Cultural Psychology* , pp. 734–759. New York: Guilford Press.

Miller, J.G. Culture and the development of everyday social explanation . *Journal of Personality and Social Psychology* , 46, 961–978. (1984) <http://dx.doi.org/10.1037/0022-3514.46.5.961>

Miller, J.G. Cultural diversity in the morality of caring: Individually-orientated versus duty-orientated impersonal codes . *Cross-Cultural Research* , 28, 3–39. (1994) <http://dx.doi.org/10.1177/106939719402800101>

Mills, J. and Clark, M.S. Exchange and communal relationships . In L. Wheeler (ed.), *Review of Personality and Social Psychology* , 3, 121–144. (1982)

Moorman, R.H. and Blakely, G.L. Individualism-collectivism as an individual difference predictor of organizational citizenship behavior . *Journal of Organizational Behavior* , 16, 127–142. (1995) <http://dx.doi.org/10.1002/job.4030160204>

Morris, M.W., Ames, D.R. and Knowles, E. (2001) What we theorize when we theorize that we theorize: Examining the 'implicit theory' constructs from a cross-cultural perspective . In G. Moskowitz (ed.), *Cognitive Social Psychology* , pp. 143–161. Mahwah, NJ: Lawrence Erlbaum Associates.

Morris, M.W. and Fu, H. How does culture influence negotiation? Dynamic constructivist analysis . *Social Cognition* , 19, 324–349. (2001) <http://dx.doi.org/10.1521/soco.19.3.324.21475>

Morris, M.W., Larrick, R.P. and Su, S.K. Misperceiving negotiation counterparts: When situational determined behaviors are attributed to personality traits . *Journal of Personality and Social Psychology* , 77, 52–67. (1999) <http://dx.doi.org/10.1037/0022-3514.77.1.52>

Morris, M.W. and Peng, K. Culture and cause: American and Chinese attributions for social and physical events . *Journal of Personality and Social Psychology* , 67, 949–971. (1994) <http://dx.doi.org/10.1037/0022-3514.67.6.949>

Murdock, P. and Provost, C. Measurement of cultural complexity . *Ethnography* , 12, 379–392. (1973)

Nelson, M.R. and Shavitt, S. Horizontal and vertical individualism and achievement values: A multimethod examination of Denmark and the United States . *Journal of Cross-Cultural Psychology* , 33, 439–458. (2002) <http://dx.doi.org/10.1177/0022022102033005001>

Nisbett, R. (2003) *The Geography of Thought* . New York: Free Press.

Nisbett, R.E. and Cohen, D. (1996) *Culture of Honor: The Psychology of Violence in the South* . Boulder, CO: Westview Press.

Nisbett, P.E., Peng, K., Choi, I. and Norenzayan, A. Culture and systems of thought: Holistic vs. analytic cognition . *Psychological Review* , 108, 291–310. (2001) <http://dx.doi.org/10.1037/0033-295X.108.2.291>

Oyserman, D. and Lee, S.W.S. Does culture influence what and how we think? Effects of priming individualism and collectivism . *Psychological Bulletin* , 134, 311–342. (2008) <http://dx.doi.org/10.1037/0033-2909.134.2.311>

Parsons, T. (1949) Essays in Sociological Theory: Pure and Applied . New York: Free Press.

Pelto, P.J. (1968) The difference between ‘tight’ and ‘loose’ societies . *Transaction* , April, 37–40.

Peng, K. and Nisbett, R.E. Culture, dialectics and reasoning about contradiction . *American Psychologist* , 54, 741–754. (1999) <http://dx.doi.org/10.1037/0003-066X.54.9.741>

Realo, A., Allik, J. and Valdi, M. The hierarchical structure of collectivism . *Journal of Research in Personality* , 31, 93–116. (1997) <http://dx.doi.org/10.1006/jrpe.1997.2170>

Redfield, R. (1941) The Folk Culture of the Yucatan . Chicago: University of Chicago Press.

Redfield, R. (1956) Peasant Society and Culture: An Anthropological Approach to Civilization . Chicago: University of Chicago Press.

Riesman, D., Glazer, N. and Denney, R. (1961) The Lonely Crowd: A Study of the Changing American Character . New Haven, CT: Yale University Press.

Sanchez Burks, J., Nisbett, R.E. and Yabarria, O. Cultural styles, relational schemas and prejudice against outgroups . *Journal of Personality and Social Psychology* , 79, 174–189. (2000) <http://dx.doi.org/10.1037/0022-3514.79.2.174>

Schwartz, S.H. (1992) Universals in the content and structure of value: Theoretical advances and empirical tests in 20 countries . In M. Zanna (ed.), Advances in

Experimental Social Psychology , 25, 1–66. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960281-6>

Schwartz, S.H. (1994) Beyond individualism and collectivism. New cultural dimensions of values . In U. Kim, ed. , H.C. Triandis, ed. , C. Kagitcibasi, ed. , S.C. Choi, ed. and G. Yoon (eds), Individualism and Collectivism: Theory, Method, and Applications , pp. 85–122. Newbury Park, CA: Sage.

Schwartz, S.H. and Bilsky, W. Toward a theory of universal structure and content of values: Extension and cross-cultural replications . Journal of Personality and Social Psychology , 58, 878–891. (1990) <http://dx.doi.org/10.1037/0022-3514.58.5.878>

Shavitt, S. et al. (2011) Horizontal an vertical individualism and collectivism . In M. Gelfand, ed. , C. Chiu, ed. and Y. Hong (eds), Advances in Culture and Psychology , pp. 309–350. New York: Oxford University Press.

Shteynberg, G., Gelfand, M.J. and Kim, K. Peering into the ‘Magnum Mysterium’: The explanatory power of descriptive norms . Journal of Cross-Cultural Psychology , 40, 46–69. (2009) <http://dx.doi.org/10.1177/0022022108326196>

Shweder, R.A. (1990) Cultural psychology — what is it? In J.W. Stigler, ed. , R. Shweder, ed. and G. Herdt (eds), Cultural Psychology , pp. 1–46. Cambridge: Cambridge University Press.

Smith, P.B., Misumi, J., Tayeb, M., Peterson, M. and Bond, M.H. On the generality of leadership style measures across culture . Journal of Occupational Psychology , 62, 97–100. (1989) <http://dx.doi.org/10.1111/j.2044-8325.1989.tb00481.x>

Smith, P.B. and Schwartz, S.H. (1997) Values . In J.W. Berry, ed. , C. Kagtchibasi, ed. , and M.H. Segall (eds), Handbook of Cross-Cultural Psychology, Vol. 3 , 2nd Edition, pp. 77–119. Boston: Allyn & Bacon.

Soh, S. and Leong, F.T.L. Validity of vertical and horizontal individualism and collectivism in Singapore . Journal of Cross-Cultural Psychology , 33, 3–15. (2002) <http://dx.doi.org/10.1177/0022022102033001001>

- Streep, P. (1995) Confucius: The Wisdom . Boston: Bullfinch.
- Toennies, F. (1957) Gemeinschaft und Gesellschaft [Community and society] . New Brunswick, NJ: Transaction Publishing.
- Torrey, E.F. and Miller, J. (2001) The Invisible Plague: The Rise of Mental Illness from 1750 to the Present . New Brunswick, NJ: Rutgers University Press.
- Tov, W. and Diener, E. (2007) Culture and subjective well-being . In S. Kitayama, ed. and D. Cohen, (eds), Handbook of Cultural Psychology , pp. 691–713. New York: Guilford Press.
- Trafimow, D., Triandis, H.C. and Goto, S.G. Some tests of the distinction between the private and the collective self . Journal of Personality and Social Psychology , 60, 649–655. (1991) <http://dx.doi.org/10.1037/0022-3514.60.5.649>
- Triandis, H.C. (1964) Cultural influences upon cognitive processes . In L. Berkowitz (ed.), Advances in Experimental Social Psychology , pp. 1–48. New York: Academic Press. <http://dx.doi.org/10.1016/S0065-2601%2808%2960047-7>
- Triandis, H.C. (1972). The Analysis of Subjective Culture . New York: Wiley.
- Triandis, H.C. (1977) Interpersonal Behavior . Monterey: Brooks/Cole.
- Triandis, H.C. (1980) Values, attitudes and interpersonal behavior . In H. Howe, ed. and M. Page (eds), Nebraska Symposium on Motivation , 1979, 27, 195–260. Lincoln, NE: University of Nebraska Press.
- Triandis, H.C. The self and social behavior in differing cultural context . Psychological Review , 96, 506–520. (1989) <http://dx.doi.org/10.1037/0033-295X.96.3.506>
- Triandis, H.C. (1994) Culture and Social Behavior . New York: McGraw-Hill.
- Triandis, H.C. (1995) Individualism and Collectivism . Boulder, CO: Westview Press.

Triandis, H.C. (2009a) Ecological determinants of cultural variations . In R.W. Wyer, ed. , C.-Y. Chiu, ed. and Y.Y. Hong (eds), Understanding Culture: Theory, Research and Applications , pp. 189–210. New York: Psychology Press.

Triandis, H.C. (2009b) Fooling Ourselves: Self-deception in Politics, Religion, and Terrorism . Westport, CT: Praeger.

Triandis, H.C., Bontempo, R., Betancourt, H., Bond, M., Leung, K., Brenes, A. and de Montmollin, G. The measurement of the etic aspects of individualism and collectivism across cultures . Australian Journal of Psychology , 38, 257–267. (1986) <http://dx.doi.org/10.1080/00049538608259013>

Triandis, H.C., Bontempo, R., Villareal, M.J., Asai, M. and Lucca, N. Individualism and collectivism: Cross-cultural perspectives on selfingroup relationships . Journal of Personality and Social Psychology , 54, 323–338. (1988) <http://dx.doi.org/10.1037/0022-3514.54.2.323>

Triandis, H.C., Chan, D.K.-S., Bhawuk, D.P.S., Iwao, S. and Sinha, J.B.P. Multimethod probes of allocentrism and idiocentrism . International Journal of Psychology , 30, 461–480. (1995) <http://dx.doi.org/10.1080/00207599508246580>

Triandis, H.C., Chen, X.-P. and Chan, D.K.-S. Scenarios for the measurement of collectivism and individualism . Journal of Cross-Cultural Psychology , 29, 275–289. (1998) <http://dx.doi.org/10.1177/0022022198292001>

Triandis, H.C. and Gelfand, M.J. Converging measurement of horizontal and vertical individualism and collectivism . Journal of Personality and Social Psychology , 74, 118–128. (1998) <http://dx.doi.org/10.1037/0022-3514.74.1.118>

Triandis, H.C., Leung, K., Villareal, M.J. and Clack, F. Allocentric vs. idiocentric tendencies: Convergent and discriminant validation . Journal of Research in Personality , 19, 395–415. (1985) <http://dx.doi.org/10.1016/0092-6566%2885%2990008-X>

Triandis, H.C., McCusker, C., Betancourt, H., Iwao, S., Leung, K., Salazar, J.M. et al. An etic-emic analysis of individualism and collectivism . *Journal of Cross-Cultural Psychology* , 24, 366–383. (1993) <http://dx.doi.org/10.1177/0022022193243006>

Triandis, H.C., McCusker, C. and Hui, C.H. Multimethod probes of individualism and collectivism . *Journal of Personality and Social Psychology* , 59, 1006–1020. (1990) <http://dx.doi.org/10.1037/0022-3514.59.5.1006>

Van de Vliert, E. Climitoeconomic roots of survival versus self-expression culture . *Journal of Cross-Cultural Psychology* , 38, 156–172. (2007) <http://dx.doi.org/10.1177/0022022106297298>

Wagner, J.A., III Studies of individualism-collectivism: Effects of cooperation in groups . *Academy of Management Journal* , 38, 152–170. (1995) <http://dx.doi.org/10.2307/256731>

Weldon, E. Deindividuation, interpersonal affect, and productivity in laboratory task groups . *Journal of Applied Social Psychology* , 14, 469–485. (1984) <http://dx.doi.org/10.1111/j.1559-1816.1984.tb02253.x>

Zemba, Y., Young, M.J. and Morris, M.W. Blaming leaders for organizational accidents: Proxy logic in collective versus individual-agency cultures . *Organizational Behavior and Human Decision Processes* , 101, 36–51. (2006) <http://dx.doi.org/10.1016/j.obhd.2006.04.007>

Zola, I.K. Culture and symptoms: An analysis of patients presenting complaints . *American Sociological Review* , 3, 615–630. (1966) <http://dx.doi.org/10.2307/2091854>

Zwier, S. (1998) Patterns of language use in individualistic and collectivistic cultures . Unpublished PhD dissertation, Free University of Amsterdam.

<http://dx.doi.org/10.4135/9781446249222.n51>

Handbook of Theories of Social Psychology: Volume 1

Theories of Social Psychology: An Introduction

Contributors: Paul A.M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Editors: Paul A. M. Van Lange & Arie W. Kruglanski & E. Tory Higgins

Book Title: Handbook of Theories of Social Psychology: Volume 1

Chapter Title: "Theories of Social Psychology: An Introduction"

Pub. Date: 2012

Access Date: January 15, 2015

Publishing Company: SAGE Publications Ltd

City: London

Print ISBN: 9781847875143

Online ISBN: 9781446249215

DOI: <http://dx.doi.org/10.4135/9781446249215.n1>

Print pages: 1-9

©2012 SAGE Publications, Inc. All Rights Reserved.

This PDF has been generated from SAGE knowledge. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781446249215.n1>

[p. 1 ↓]

Theories of Social Psychology: An Introduction

Paul A.M. Van Lange, ed. , Arie W. Kruglanksi, ed. , and E. Tory Higgins, ed.

The advancement of theory is a key goal of science. Ideally, theory helps us explain particular events and phenomena and find the underlying *truth, beyond surface appearances*. Theory helps us see the coherent structures in seemingly chaotic environments and make inroads into previously uncharted domains, thus affording progress in the way we understand the world around us. Because it elucidates the causal mechanisms that produce manifest effects, theory points to ways of intervening in phenomena and changing the course of events; hence, theory is of essential pragmatic value and constitutes an indispensable tool for application.

This book is about theories of social psychology, a field broadly defined by reciprocal influences between individuals and their social environments – that is, other people, whether present, imagined, or implied (Allport, 1954; cf. Deutsch and Krauss, 1965: 1; Jones and Gerard, 1967: 1; Shaver, 1977: 4; Van Lange, 2006: 13). Consistent with this definition, social psychology encompasses a wide range of domains addressing the manifold targets of influence (including individuals' thoughts, feelings, and actions), as well as different kinds of influence (conscious and unconscious; implicit and explicit). Because of its subject matter, social psychology is critically relevant to a vast multitude of events in the social realm spanning the gamut from intraindividual judgments and decisions to problems in interpersonal relations, group and intergroup dynamics all the way to the impacts of culture and intercultural encounters. Indeed, social psychologists have been carrying out important conceptual and empirical work in all these domains of endeavor.

Given the scope and relevance of social psychology, and the importance of theory for a field of science, it is somewhat surprising that considerable time has passed since a

volume was published specifically devoted to social psychological theories. The latest effort of this kind was the 1980 primer by West and Wicklund, which complemented the 1970 book by Shaw and Costanzo (revised in 1982), and the earlier classic work by Deutsch and Krauss (1965). Yet considerable theoretical work in social psychology has been accomplished in the three decades that followed. Numerous conceptual frameworks have been constructed at the various levels of social psychological analysis. Moreover, [p. 2 ↓] theoretical debates and commentaries have appeared on approaches to social theory construction (e.g., Higgins, 2004; Kelley, 2000; Kruglanski and Higgins, 2004), the state of social/personality theory (Kruglanski, 2001; Mischel, 2004), and the theoretical bridging of social psychological analyses with other fields and scientific disciplines (Kruglanski, 2006; Van Lange, 2006, 2007). The *Handbook of Theories of Social Psychology* reflects these discussions and provides a comprehensive perspective on major theoretical developments since the time that social psychology was still in its infancy. As such, it depicts our discipline's journey from childhood to adulthood, spanning more than half a century of growth and conceptual progress in scientific understanding of the social world.

Theory: Regulatory Ideals

It is not easy to precisely define what a theory is – and more importantly, what would qualify as theory and what would not. Like earlier writers, we suggest that a theory may minimally be defined as a set of interrelated propositions (or principles) concerning a phenomenon or set of phenomena (Mandler and Kessen, 1959: 159; Shaw and Costanzo, 1982: 4). Clearly, theories may differ in their generality, precision, and origins. Our approach in this volume has been inclusive rather than exclusive, and based on a minimalist definition of the concept of “theory” to guide our final decision. We reasoned that a conceptual framework that inspired significant empirical research is worthy of inclusion even if it is incomplete or otherwise imperfect from a “purist” metatheoretical perspective. Besides, social psychological theories tend to be “middle range” (Merton, 1949: 5) anyway and hence “intermediate to minor working hypotheses” rather than grand theoretical edifices. Social psychology is rich in such mid-range theories, each representing a “work in progress” rather than being an etched-in-stone master scheme (see Pinker, 2002: 241; Van Lange, 2006: 8).

Whatever definition of a theory one opts for, it is important to ascertain what constitutes a good theory. Though numerous constructs have been advanced to outline various qualifications, standards, and criteria for theoretical “goodness,” there is a fair amount of consensus at least concerning these matters: theories are believed to be better if they have greater explanatory power; are more suitable to empirical tests and modeling; are more “logical,” in the sense of coherence and internal consistency; are capable of explaining more (phenomena) with less (by way of assumptions) reflecting the criterion of parsimony or Occam’s Razor; and, critically, inspire new research that yields empirical discoveries (see, for example, Fiske, 2004; Higgins, 2004). For the present purposes, we focus on a framework that is characterized by four broad regulatory ideals, namely: truth, abstraction, progress, and applicability (see also Kruglanski, 2006). We discuss each of these in turn.

Ideal 1: Truth

A theory should be dealing with the truth; it should separate fact from fiction; it should establish what’s real rather than what’s imaginary. Although an inaccurate, fictitious theory can serve important functions (such as serving an heuristic function to stimulate further research), it should be clear that theories seek to pursue “the truth and nothing but the truth.” This is what hypothesis testing is all about. The entire logic of experimental design is to eliminate (or prove invalid) possible alternative interpretations of empirical facts. Critical experiments are designed to set apart competing theories and decide which one appears more valid, and is better supported by the available evidence, than its competitors. And holding onto a theory known to be false simply because it had heuristic or communicative value (i.e., it was [p. 3 ↓] easy to convey to others) is self-contradictory because subscribing to a theory is tantamount to believing it to be true.

Yet, as a regulatory ideal, truth can be striven for but never securely attained. No theory, however successful, is secure, for alternative accounts of the same data are always possible in the future even if they may not be apparent in the present. You can disprove a theory but never prove it. You can only find support for a theory based on what is currently known. Moreover, the empirical “facts” are far from absolute. As noted by Popper (1959: 111) the empirical basis of science is both conjectural and fallible:

The empirical basis of objective science has nothing “absolute” about it. Science does not rest upon rock bottom. The bold structure of its theories rises, as it were, above a swamp. It is like a building erected upon piles. The piles are driven down from above into the swamp; and when we cease our attempt to drive our piles into a deeper layer, it is not because we have reached firm ground. We simply stop when we are satisfied that they are firm enough to carry the structure, at least for the time being.

Thus, although a primary ideal of science, the pursuit of truth is a mission that can never be completely accomplished.

Ideal 2: Abstraction

A theory should be the result of abstraction, in that the particulars (e.g., phenomena, events) need to be described in terms of the general (concepts, assumptions, principles). While a particular phenomenon may be interesting in and of itself, one needs a theory to understand the psychological principles that underlie the phenomenon – the same principles that underlie other seemingly disparate phenomena as well. It is the higher level of aggregation that a theory should pursue, to transcend particular observations and link them at a deeper (i.e., more abstract) level to other observations. Thus, theories focus on the heart of the matter in terms of understanding and insight, as it deals with essential causal mechanisms underlying observed effects.

Ideal 3: Progress

Any new theory is expected to make a contribution beyond what was previously known; it should improve or expand our explication of a given realm of phenomena representing the ideal of progress. It should replace myths by wisdom, and it should add truth to existing knowledge enlarging the scope of our understanding. Ideally then, newer theories relate to and build upon past theories, replacing inaccurate with accurate principles, or complementing a predecessor theory with new principles that had not previously been identified. Science is unlikely to progress if theories are not subject

to refinement through a process of sharpening and empirical testing. In this sense, the principle of progress is closely linked with the principle of truth, as the theoretical refinements and modifications are in the service of ever-greater validity and precision. Because truth is an important regulatory (though largely unattainable) goal of a theory, a theory is often subject to refinement and precision; for example, by outlining the conditions in which the hypotheses derived from the theory should be supported. Also, a theory often inspires new ways of thinking, because it serves (implicitly, at least) as a tool for theorists and researchers to see connections and relationships that would not have been evident on the basis of data alone (cf. Shaw and Costanzo, 1982). Finally, a theory is often an inspiration for new research questions, along with new tools, methodologies, and paradigms (Fiedler, 2004; Fiske, 2004). As such, theories function both as bridges to the past (the past findings a theory accounts for) and the future (future research and findings inspired by the theory). Because the implications of a theory inspire new predictions that in turn inspire new research to test them, a theory is the driving force behind new empirical discoveries about what the [p. 4 ↓] world is and how it works – progress from new discoveries (Higgins, 2004).

Ideal 4: Applicability

Ideally, a psychological theory should speak to many events and issues in everyday life. It should be applicable to real-world concerns and afford interventions aimed to alter the course of events in desirable ways. As Edward E. Jones (1986: 100) aptly remarked, “The future of social psychology is assured not only by the vital importance of its subject matter but also by its unique conceptual and methodological strengths that permit the identification of underlying processes in *everyday social life* [italics added]. Just as scientific progress is closely linked to the quest for truth, a theory’s applicability is closely linked to the precept of abstraction. In other words, the more abstract the theory, the greater its empirical content (Popper, 1959) and the broader the range of situations to which it applies. Of course, theoretical breadth in and of itself is not tantamount to application, and an appreciable measure of ingenuity is needed to translate a theory’s implications into specific procedures and interventions of practical value.

In fact, despite the intimate relation between theory and application, the two have been often juxtaposed with each other, and viewed as fundamentally disparate – theory

versus practice. Theory has been often associated with logic, deduction, and knowledge ('knowing'), whereas application has been often associated with intuition, induction, and implementation ('doing'). Perhaps Kurt Lewin's famous dictum, "Nothing is as practical as a good theory," received so much attention because it was surprising in light of the general tendency to view application as the very antithesis of theorizing. Nonetheless, the notion of "translational research" highlights the intimate connection between theory and application and encourages theoreticians (often by means of funding opportunities) to descend from the "Olympus of pure thought" and explore the possible contribution of their ideas to the solution of the multitude of real-world problems to which they pertain.

Theories Need to Be Tapas-Proof

We have conceptualized theory construction in terms of four regulatory ideals: namely, Truth, Abstraction, Progress, and Applicability. These may serve as Standards for critically evaluating a theory (TAPAS). Implicitly, these TAPAS are used pervasively for assessing psychological research. Truth claims are evaluated via appropriate experimental design, the carrying out of critical experiments, and by conceptual reviews and meta-analyses of the available evidence. Abstraction is often used to evaluate theoretical breadth and, at a more concrete level, in testing general psychological hypotheses via specific empirical observations. Progress is assessed through the ubiquitous demand that research be innovative and makes a significant new contribution to knowledge. Finally, applicability is often intimated by the plea for a body of psychological work to have "broader impact," (e.g., Buunk, 2006; Fiedler, 2006; Van Lange, 2006) and it is deemed an important criterion for the funding of research grants at federal granting agencies in the US (like the National Science Foundation and the National Institutes of Health) because it is understood that scientific knowledge should have societal benefits. We highlight TAPAS explicitly in the hope that this will enhance their accessibility and use among social psychological theorists.

Theory Construction and Development

Where do ideas come from? What are the sources of theorists' inspirations? How do social psychologists turn tacit hunches and inchoate intuitions into overtly articulated [p. 5 ↓] theoretical statements? And how does one facilitate theoretical progress? How "responsible" is the theorist for the future fate of her or his intellectual offspring? There is little public discussion of these matters in the social psychological literature though experience on these matters has likely accumulated in the personal memory stores of individual theorists. The latter assumption served as a departure point for a special issue published in the *Personality and Social Psychology Review* (2004, Vol. 8), entitled "Theory construction in social personality theory: Personal experiences and lessons learned." Contributors to this special issue shared with the readers their innermost insights and metatheoretical self-reflections as well as "tricks of the trade" and personal strategies of theory construction and development. For example, there was a discussion about how detecting features that different phenomena share in common can lead to the development of overarching theories and the unveiling of deeper structures masked by striking surface differences (e.g., Kruglanski, 2004). Another paper discussed ways of carrying out theoretically inspired research programs and ensuring that a theory's potential for contribution and impact is realized to its fullest (Higgins, 2004). Yet other papers discussed how to benefit from sabbatical leaves by developing a broader perspective on a body of empirical work (Zanna, 2004), and how collaboration with one's colleagues may deepen and enrich one's theoretical frameworks (Levine and Moreland, 2004). The chapters in the present volume illustrate such fascinating strategies and approaches via numerous theorists' personal narratives of their intellectual journeys; journeys wherein their conceptual insights were explicitly articulated, and their theoretical structures were erected and developed.

Imparting Theory Construction

It has been argued that over the course of its history, social psychology's attention to theory has declined, while the field has become increasingly data-driven and phenomenon-focused (e.g., Fiedler, 2004; Kruglanski, 2001; see also Jones, 1986).

Part of the problem might have been the lack of systematic ways of teaching theory construction and incorporating courses and workshops on this topic in graduate curricula in social psychology. To a considerable extent, our graduate training focuses on issues of method, design, and data analysis. Theory construction is generally treated as unteachable and largely a matter of inspiration. Yet a great deal about theorizing can be articulated and imparted. The important ingredients of a successful theory, what we called TAPAS, can be defined, explained, and deliberately striven for. And the way those theoretical properties have been cultivated and developed by various theorists has also been explicitly described (Kruglanski and Higgins, 2004). In this vein, the present volume too purports to afford greater appreciation of the theorizing enterprise through personal stories of successful theoreticians, thereby teaching theoretical skills through examples.

Admittedly, passive exposure to historical accounts of theorizing strategies is unlikely to suffice. Concrete practice in translating the general principles of theory construction into actual attempts at conceptualization seems essential. One of us has briefly described a seminar course that aims to do this (Higgins, 2004). And regarding the teaching of applicability, a recent book by Buunk and Van Vugt (2007) provides a platform for practicing the skills of making theoretical concepts applicable to specific real-world problems. It does so by challenging the student to formalize the key properties of an important social issue (e.g., how to reduce vandalism in football stadiums; how to increase environmental concerns), analyze it in cause-and-effect terms using social psychological concepts and principles, and suggest possible policy measures based on such theoretical analysis of the problem.

[p. 6 ↓]

About This Handbook

Like the earlier books on social psychological theories, the present handbook reviews major conceptual developments in our field. Yet it also differs from prior reviews in significant respects. To begin with, the handbook covers the decades that passed since the last such volumes appeared in print (namely, West and Wicklund, 1980; Shaw and Costanzo, 1982) during which significant conceptual developments have

taken place. Within that period social psychology has undergone exponential growth worldwide, yielding an explosion of theoretical frameworks by social psychologists at all levels of analysis – the biological system, the cognitive system, the motivational (and affective) system, the interpersonal system, and the group and cultural system (cf. Higgins and Kruglanski, 1996; Kruglanski and Higgins, 2007). At the same time, classic theories continued to inspire research in their different incarnations, such as Festinger's dissonance theory (1957) that received a "new look" in Cooper and Fazio's interpretation, and theories that have undergone a transition, such as Bandura's social learning theory that shifted into his social cognitive theory (for more information, see Chapters by Cooper and Bandura in this handbook).

In addition to offering a more up-to-date portrait of the theoretical landscape in social psychology, this handbook complements its predecessors in at least three important ways. First, each author has provided a personal, historical narrative of the theory's development, including the various inspirations, serendipitous events, critical junctures, and problem-solving efforts that affected theoretical choices and influenced the theory's evolution and impact. These personal narratives are unique to the present handbook and provide a richly textured background for understanding how theories are actually created, nurtured, and shaped over time.

Second, each author has placed her or his theory within the intellectual history of the topic it addresses, and has commented on the theory's unique contribution to the field against this intellectual backdrop. This places each theory within a second kind of history – *the history of ideas* – that is also unique to this handbook. This aspect of the book provides a strong answer to the question, "Why should I care? What's the added value gained from the theory's contribution?" Third, each author has evaluated her or his theory in terms of its *applicability* for not only understanding but also for solving critical social issues and real-world problems. This aspect of the book provides a second strong answer to the question, "Why should I care? What's the added value?"

The overarching principle underlying these three foci was that theories are ultimately about *ideas*. Hence, it is important to learn: (a) *where* the ideas came from and how they developed; (b) *why* the ideas matter intellectually and historically; and (c) *what* difference the ideas make for dealing with current societal concerns.

Criteria for Inclusion

When is a conceptual framework a theory, when is it a model, and when is it a hypothesis? Though presumably social psychologists would agree on the theoretical “status” of numerous conceptual frameworks in the field, they might disagree with respect to some. We do not regard this surprising because the field of social psychology has grown considerably – and in many directions – over the past three decades (since the 1982 book by Shaw and Costanzo). Further, social psychology lacks a widely received “grand theory” of the sort that seeks to account for all diversity and serves as the consensual platform for more specialized analyses (like evolutionary theory in biology or rational choice theory in economics).

In selecting theories for the present handbook, we used the following broad guidelines. First, we decided to include [p. 7 ↓] theories with a bit of a “track record”; theories whose development could be traced back more than a decade. Indeed, one of our objectives was to provide insights into the ways that social psychologists design, develop, and “nurture” their theories. And this process, inevitably, takes time. It takes time to develop a theory’s validity (the *truth* criterion), a theory’s generality (the *abstraction* criterion), a theory’s generative power (the *progress* criterion), and a theory’s usefulness in solving real-world problems (the *applicability* criterion).

Second, we decided to include theories that have survived the passage of time, those that continue to guide research in the “here and now.” It is, of course, possible that some theories will be revitalized in some way in the future, in which case they may well be included in future handbooks on theory.

Finally, we decided to include theories developed in the tradition of social psychology, rather than ones developed outside of the field of social psychology. Admittedly, there are many influential theories and models on social psychological topics that could be included. Examples are theories in fields such as cognitive neuroscience, decision-making, economics, sociology, and political science. But to include such theories, we felt, would excessively broaden the scope of our enterprise and defocus our perspective from conceptual work carried out in social psychology proper.

We should note that not all theories that met these guidelines ended up in this book. In some cases, the author of a theory had passed away. In some cases, the author of a theory declined our invitation. We regret not having these theories represented. But we are very thankful for the theories that are represented, and we are extremely grateful to our authors for their contributions to this Handbook. It has been a wonderful experience for us as Editors to learn more about the history of their landmark theories.

References

- Allport, G.W. (1954) The historical background of modern social psychology . In G. Lindzey (ed.) *Handbook of Social Psychology* , 1, 3–56. Cambridge, MA: Addison-Wesley.
- Buunk, A.P. (2006) Social psychology deserves better: Marketing the pivotal social science . In P.A.M. Van Lange (ed.), *Bridging Social Psychology-Benefits of Transdisciplinary Approaches* , pp. 83–89. Mahwah, NJ: Erlbaum.
- Buunk, A.P. and Van Vugt, M. (2007) *Applying Social Psychology: From Problems to Solutions* . Thousand Oaks, CA: Sage.
- Deutsch, M. and Krauss, R.M. (1965) *Theories in Social Psychology* New York: Basic Books.
- Einstein, A. (1934) *Mein Weltbild* . Amsterdam: Querido.
- Festinger, L. (1957) *A Theory of Cognitive Dissonance* . Stanford, CA: Stanford University Press.
- Fiedler, K. (2004) Tools, toys, truisms, and theories: Some thoughts on the creative cycle of theory formation . *Personality and Social Psychology Review* , 8, 123–131.
http://dx.doi.org/10.1207/s15327957pspr0802_5
- Fiedler, K. (2006) On theories and societal practice: Getting rid of a myth . In P.A.M. Van Lange (ed.), *Bridging Social Psychology: Benefits of Trans-disciplinary Approaches* , pp. 65–70. Mahwah NJ: Erlbaum.

Fiske, S.T. (2004) Mind the gap: In praise of informal sources of formal theory . *Personality and Social Psychology Review* , 8, 138–145. http://dx.doi.org/10.1207/s15327957pspr0802_6

Higgins, E.T. (2004) Making a theory useful: Lessons handed down . *Personality and Social Psychology Review* , 8, 138–145. http://dx.doi.org/10.1207/s15327957pspr0802_7

Higgins, E.T., ed. and Kruglanski, A.W. (eds) (1996) *Social Psychology: Handbook of Basic Principles* . New York: Guilford Press.

Jones, E.E. (1986) Major developments in social psychology during the past five decades . In G. Lindzey, ed. and E. Aronson (eds), *The Handbook of Social Psychology* , 3rd Edition, pp. 47–135. New York: McGraw-Hill.

Jones, E.E. (1988) Major developments in five decades of social psychology . In D.L. Gilbert, ed. , Susan T. Fiske, ed. and G. Lindzey, ed. , *The Handbook of Social Psychology* , 1, 3–57. New York: McGraw-Hill.

Jones, E.E. and Gerard, H.B. (1967) *Foundations of Social Psychology* . New York: Wiley.

Kelley, H.H. (2000) The proper study of social psychology . *Social Psychology Quarterly* 63, 3–15. <http://dx.doi.org/10.2307/2695877>

Kruglanski, A.W. (2001) That ‘vision thing’: The state of theory in social and personality psychology at the edge of the new millennium . *Journal of Personality and Social Psychology* 80, 871–875. <http://dx.doi.org/10.1037/0022-3514.80.6.871>

Kruglanksi, A.W. (2004) The quest for the gist: On challenges of going abstract in social and personality psychology . *Personality and Social Psychology Review* , 8, 156–163. http://dx.doi.org/10.1207/s15327957pspr0802_9

Kruglanski, A.W. (2006) Theories as bridges . In P.A.M. Van Lange (ed.), *Bridging Social Psychology: Benefits of Transdisciplinary Approaches* , pp. 21–34. Mahwah, NJ: Erlbaum.

Kruglanski, A.W. and Higgins, E.T. (2004) Theory construction in social and personality psychology: Personal experiences and lessons learned . *Personality and Social Psychology Review* , 8, 96–97. http://dx.doi.org/10.1207/s15327957pspr0802_1

Kruglanski, A.W., ed. and Higgins, E.T. (eds) (2007) *Social Psychology: Handbook of Basic Principles* , 2nd Edition. New York: Guilford Press.

Levine, J.M. and Moreland, R.L. (2004) Collaboration: The social context of theory development . *Personality and Social Psychology Review* , 8, 164–172. http://dx.doi.org/10.1207/s15327957pspr0802_10

Mandler, G. and Kessen, W. (1959) *The Language of Psychology* New York: Wiley.

Merton, R.K. (1949) *Social Theory and Social Structure* . Glencoe, IL: The Free Press.

Mischet, W. (2004) Toward an integrative science of the person . *Annual Review of Psychology* 55, 1–22. <http://dx.doi.org/10.1146/annurev.psych.55.042902.130709>

Pinker, S. (2002) *The Blank Slate: The Modern Denial of Human Nature* . New York: Viking.

Popper, K.R. (1959) *The Logic of Scientific Discovery* . New York: Harper. (Original work published as Logikder Forschung, 1935).

Shaver, K.G. (1977) *Principles of Social Psychology* . Cambridge, MA: Winthrop.

Shaw, M.E. and Costanzo, P.R. (1982) *Theories of Social Psychology* . New York: McGraw-Hill. (Originally published 1970).

Van Lange, P.A.M. (ed.) (2006) *Bridging Social Psychology: Benefits of Transdisciplinary Approaches* . Mahwah, NJ: Erlbaum.

Van Lange, P.A.M. (2007) Benefits of bridging social psychology and economics . *Academy of Management Review* , 32, 671–674. <http://dx.doi.org/10.5465/AMR.2007.24352124>

West, S.G. and Wicklund, R.A. (1980) A Primer of Social Psychological Theories .
Monterey, CA: Brooks/Cole.

Zanna, M. (2004) A naïve epistemology of a working social psychologist (Or the working epistemology of a social psychologist): The value of taking “temporary givens” seriously . *Personality and Social Psychology Review* , 8, 210–218. http://dx.doi.org/10.1207/s15327957pspr0802_15

<http://dx.doi.org/10.4135/9781446249215.n1>