

About the Authors

Robert R. McCrae, PhD, is Research Psychologist at the Gerontology Research Center of the National Institute on Aging. He received his doctorate in personality psychology from Boston University in 1976, and has since conducted research on personality structure, assessment, and development. His recent work has centered on cross-cultural studies of personality. He has authored or coauthored over 250 articles and chapters, and with Paul T. Costa, Jr., he is author of the Revised NEO Personality Inventory.

Paul T. Costa, Jr., PhD, is Chief of the Laboratory of Personality and Cognition at the National Institute on Aging's Intramural Research Program in Baltimore, Maryland. His enduring interests are in the structure and measurement of personality and in lifespan development, psychopathology, and neurobiological bases of personality. He has authored or coauthored over 300 papers and chapters and has served as President of APA Divisions 5 and 20, the Association for Research in Personality, and the International Society for the Study of Individual Differences.

The longitudinal studies of personality on which this book is chiefly based have resulted from the collaborative efforts of a number of distinguished researchers and thousands of dedicated volunteers. This book has benefited most directly from the generosity and commitment of the men and women participants in the Baltimore Longitudinal Study of Aging. We dedicate it to them, and to Bill and Karol.

Preface

During the 1970s several major longitudinal studies of personality were completed and published, providing for the first time a solid scientific basis for evaluating theories of personality and aging. To almost everyone's surprise, all these studies pointed to an extraordinary degree of stability: Personality apparently changes little after age 30 in most people. After several years of testing our interpretation of the data and working out some of its implications for personality psychology and for an understanding of adult life, we published a small book on the topic (McCrae & Costa, 1984).

In the next few years several additional longitudinal studies were reported that generally confirmed early findings. But the major development of the 1980s was the rediscovery of the Five-Factor Model (FFM) and the demonstration that it was indeed a comprehensive model of personality trait structure, including virtually all the traits identified in common language and in scientific theories. This was arguably the most important advance in modern personality psychology, because it put to rest the unfruitful competition between rival trait models that had gone on for decades, and it allowed a systematic approach to the study of personality. We brought out a new version of our book in 1990 that described the FFM and its implications for the study of aging.

The current revision of our book *Personality in Adulthood* reports progress over the past decade. There are three notable changes since the last edition.

First, detailed research in large samples has made it possible (and necessary) to qualify our original claims about stability. Although it re-

mains the case that stability is a very good first approximation to the truth in describing the course of personality traits in adults over age 30, it is also clear that there are exceptions and qualifications: small but replicable declines in some traits, major changes in some individuals (such as those suffering from Alzheimer's disease), and a decay of the stability of individual differences over long periods of time.

The second major change, reflected in a new chapter (5), was made possible by the worldwide acceptance of the FFM. Translations of the Revised NEO Personality Inventory into over 40 languages have been used to address questions of the cross-cultural generalizability of developmental trends. Although only cross-sectional studies have so far been reported, they do much to further our understanding of personality stability and change.

Finally, another new chapter (10) reports a conceptual advance. Five-Factor Theory (FFT) is a description of the personality system, positioning traits in the broader context of the person and the world. Initially developed to explain how traits could remain stable as individuals continued to adapt to a changing world, FFT is supported by cross-cultural, comparative, and behavior genetic evidence. Concepts from the theory are used in the last chapter to structure a discussion of the effects of personality on the life course and self-concept. The subtitle of this edition of the book reflects the importance we assign to this theoretical perspective.

Although the main message in our book has endured, a wealth of new findings and new interpretations have emerged since it was first published. We hope this revised edition gives a sense of the progress the field has made.

Facts and Theories of Adult Development

Suppose for a moment that people consulted lifespan developmental psychologists as they do fortune tellers—to get a glimpse of their own future. What would we tell them to expect as they grow older? Are there predictable crises ahead? Are they likely to continue to mature and grow, or is it all downhill from here? Will their basic natures and temperaments remain essentially as they are, or will internal unfolding or changing circumstances (such as wars, illness, or technological innovations) reshape existing personalities? Do married couples grow apart with the years, or do they come to resemble each other in personality as they sometimes seem to do in appearance?

If asked to make these kinds of predictions for individuals, we would hedge—and properly so. We would point out that the scientific study of adulthood is young and little is known with certainty. Only in the past three decades have a substantial number of investigators been active in the field, and these have succeeded mainly in framing useful questions, not in providing definitive answers to them. We would also emphasize that science is concerned with generalizations, not specifics. Epidemiologists, for example, can tell us the life expectancy of the average man or the average woman and some of the factors (such as smoking and exercise) that influence longevity, but they certainly cannot predict the exact age of death for any particular individual. Too many people smoke and drink and live to 90 and too many athletes die young to allow anything more than statements of probability.

But individuals inevitably apply these statements to themselves. When Gail Sheehy published *Passages* (1976)—and later *New Passages* (1996)—millions read the book, not because of a disinterested curiosity about human development, nor because they admired her prose style (engaging as it was). People read *Passages* because they wanted to make sense of their own past, present, or future lives. In short, the topic of this book is likely to be of personal as well as academic interest to most readers, and our approach must take that fact into consideration. We will argue for a particular position fully aware that many people find it unappealing. We will therefore try to anticipate objections, and in general we will adopt an approach that Salvatore Maddi (1976) characterized as “partisan zealotry” rather than “benevolent eclecticism” in order to “provide the reader with a vivid account” (p. 2) of our views. We believe that we can accomplish that goal without sacrificing scientific objectivity, and we hope our presentation will stimulate lively discussion and further research.

The first version of this book (McCrae & Costa, 1984) took a simple but radical position: We argued that personality was stable in adulthood—that the traits one showed at age 30 would remain essentially unchanged into old age. Much more information from longitudinal, cross-sectional, and cross-cultural studies is now available, and argument has become more nuanced. Newer studies confirm that stability is the predominant feature of personality in adulthood, but they also document predictable changes at certain ages and in certain individuals. The story has become a bit more interesting.

Some readers are likely to have a strong background in personality psychology but less knowledge of gerontology; some the reverse. We will try to accommodate both groups by reviewing some fundamentals in each discipline. Although our conclusions are driven by data and our research has been squarely in the tradition of quantitative empiricism, we will not burden the reader with much technical detail about the studies we discuss—the cited literature can be consulted for that. We will, however, spend considerable time on the logic of research, specifically how scientific questions should be formulated and how particular measures, samples, or analyses can be used to answer them. Because aging is a relatively new field and personality psychology a contentious one, there are a large number of issues to address. We will consider the problems of distinguishing aging from generational and time-of-measurement effects, the validity of self-report methods of assessment,

the adequacy of a trait theory of personality, and the advantages and dangers of interviews as a source of data on personality. At each step we will try to weigh the evidence carefully, taking into account both strengths and limitations. We can state our point of view in a few paragraphs, but a critical examination of it will require the whole book.

Over the past 20 years there have been remarkable changes in the science of personality. Trait psychology, often considered passé in the 1970s, has come back with a vengeance and is now the dominant paradigm in personality psychology. We know much more about the origins and influences of traits than we did—a new understanding that has been deeply intertwined with our discoveries about lifespan development. Perhaps most crucially, we now know the scope of personality traits. The Five-Factor Model (FFM; Digman, 1990; McCrae & John, 1992) has been widely accepted as an adequate taxonomy of personality traits, and literature reviews are now routinely organized by classifying measures along the lines of these five factors. These factors—Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness—recur throughout this book; they are the basic dispositions that, as we will see, endure through adulthood and help to shape emerging lives. The factors and some representative traits that define them are listed in Table 1.

The Pendulum of Opinion on Personality Stability

When psychologists first asked themselves what happens to personality across the lifespan, they found a great deal to say about infancy, childhood, and adolescence. Most assumed, however, that adulthood was the endpoint of personality development (an *adult*, the dictionary tells us, is a fully developed individual). William James (1890), in a now famous dictum, claimed that by age 30 character was “set like plaster.” Sigmund Freud wrote volumes on the first few years of life but almost nothing on the later years; certainly they held no major role in his theory of personality. The parallel to other forms of development seemed obvious: By age 20 the vast majority of men and women have reached their full height, and—although they may settle a bit over the years—the tall remain tall, the short, short. The same seems to be true for certain kinds of intelligence. Why should we expect anything different in the case of emotionality or warmth or modesty?

It was therefore a feat of great intellectual daring to propose that

TABLE 1. The Five-Factor Model of Personality

Neuroticism	Agreeableness
Calm—Worrying	Ruthless—Softhearted
Even-tempered—Temperamental	Suspicious—Trusting
Self-satisfied—Self-pitying	Stingy—Generous
Comfortable—Self-conscious	Antagonistic—Acquiescent
Unemotional—Emotional	Critical—Lenient
Hardy—Vulnerable	Irritable—Good-natured
Extraversion	Conscientiousness
Reserved—Affectionate	Negligent—Conscientious
Loner—Joiner	Lazy—Hardworking
Quiet—Talkative	Disorganized—Well-organized
Passive—Active	Late—Punctual
Sober—Fun-loving	Aimless—Ambitious
Unfeeling—Passionate	Quitting—Persevering
Openness to Experience	
Down-to-earth—Imaginative	
Uncreative—Creative	
Conventional—Original	
Prefer routine—Prefer variety	
Uncurious—Curious	
Conservative—Liberal	

Note. Adapted from Costa and McCrae (1986c).

psychological development might continue throughout life, and one of history's boldest thinkers, Carl G. Jung, was among the earliest to take this step. His chapter on "The Stages of Life" in *Modern Man in Search of a Soul* (Jung, 1933) foreshadowed many of the central ideas in gerontological thinking, including the curve of life (Bühler, 1935), the rise of the repressed (Levinson, Darrow, Klein, Levinson, & McKee, 1978), the feminization of men and masculinization of women (Gutmann, 1970), disengagement theory (Cumming & Henry, 1961), and the midlife crisis (Jacques, 1965).

A more elaborated and systematic position was offered by Erik Erikson (1950), who postulated stages of psychosocial development to parallel Freud's stages of psychosexual development, then extended them beyond adolescence and across the remaining years of life.

The next few decades saw the beginnings of empirical research on personality and aging, some of it guided by the theories of Erikson or Jung, much of it in search of new theoretical perspectives (J. Block,

1971; Butler, 1963; Lowenthal, Thurner, & Chiriboga, 1975; Maas & Kuypers, 1974; Neugarten, 1964; Reichard, Livson, & Peterson, 1962). In the late 1970s, a new generation of theories of adult development emerged (Gould, 1978; Levinson et al., 1978; Vaillant, 1977), Sheehy's *Passages* (1976) became a major best-seller, and the popular press began to feature stories on crises in adulthood, particularly the midlife crisis.

During the 1970s there was also a proliferation of undergraduate and graduate programs in human development and gerontology. Most of these programs were explicitly interdisciplinary, examining the sociology, biology, and economy of aging as well as its psychology. Stage theories of adult development had a powerful appeal as a way of integrating such diverse material: Predictable changes in personality might prepare the individual for the social transitions and economic changes of adult life.

All these intellectual developments were consistent with the *zeitgeist*. The 1950s had focused on children; the 1960s, on youth. As the baby boomers passed 30, their lives still seemed to form the center of the culture's interest. Personal growth and development were promised by humanistic psychology, and theories of life stages seemed to fill a particular need. Personal problems could be attributed to universal developmental changes; predictable crises offered both security and spice to adult life.

The same period also saw the "graying of America," a dramatic increase in the proportion of men and women living beyond age 65, and a concomitant increase in their awareness of their economic and political power. Older people began to demand attention, and academics took up the challenge. Personality development offered an attractive alternative to studies of cognition, where decline, if not inevitable, was the general rule (Arenberg & Robertson-Tchabo, 1977; Salthouse, 1989).

But it is the nature of science to be self-correcting. Not only did scientific ideas generate theories of adult development; they also led to research. Instead of talking about what might be or about what we want to be, we can use the research efforts of the past 40 years to see what is really going on. More and more, we believe, the findings are coalescing into a pattern, and the pattern is one of predominant stability (Costa & McCrae, 1980c). Maddox (1968) showed that well-adjusted elderly people remained active. Havighurst, McDonald, Maculen, and Mazel (1979), who studied professional careers, were led to formulate what they called "continuity theory." Neugarten (1982) propounded the no-

tion of *age irrelevance* to account for the fact that age is not a very useful predictor of social functioning. As we shall see later, within the field of personality research this emphasis on stability has been strongly seconded by the work of investigators like Jack Block (1981) and Ilene C. Siegler and colleagues (see Costa, Herbst, McCrae, & Siegler, 2000). It is beginning to appear as if James and Freud were right.

But if nothing happens with age, why write a book?

This reaction is shared by some of our colleagues in the field, and it is one with which we have often confronted ourselves. One answer would be to persuade people about stability, to disillusion those who are looking for some magic transformation with age or to reassure those who fear that they face periods of developmental crisis and turmoil and would much prefer to continue the business of their lives.

There is also another, better answer. We have not said, nor will we say, that nothing of psychological importance occurs in adulthood. People live most of their lives in this period; they begin careers, raise children, fight wars, and make peace; they experience triumph and despair, boredom and love. Old age, too, has its share of new experience and new perspectives on old experience. All of this makes a fascinating story (Gullette, 1989). From our point of view, it is all the more fascinating since one of the keys to the story is the individual's personality. People stay much the same in their basic dispositions, but these enduring traits lead them to particular and ever-changing lives.

In Search of a Phenomenon

Most sciences start with a phenomenon and try to explain it. Astronomy arose from attempts to account for the regular changes observed in the moon and stars. Biology tries to explain how different species have come to exist and adapt so differently to their environments. Cultural anthropology began in efforts to explain the puzzling customs of preliterate societies, just as abnormal psychology developed from observation of the bizarre behavior of the mentally ill. But if we ask what students of aging and personality are trying to explain, we are likely to draw a blank. The field of adult personality development seems to have emerged as an afterthought, a logical extension of other branches of study.

Some researchers came to it by way of gerontology, the study of aging. We know that there are major changes in physiology with age, and the popular belief that old people begin to lose their memory has been

confirmed by controlled longitudinal studies that demonstrate declines in certain, though not all, cognitive abilities (Salthouse, 1989). By analogy, some investigators began to wonder about personality. Does it too change with age? Is there a gradual decline in emotional or social functioning? Do older people become increasingly susceptible to mental illness, as they do to physical illness?

Researchers who began as students of personality had a somewhat different basis for their questions. We know (or think we know) that there are changes in personality in childhood and adolescence. Infants become emotionally responsive to familiar faces only around 30 days; at 8 months they are likely to develop separation anxiety when taken away from their parents. Middle childhood is a period of compliance for most children; adolescence is generally conceded to be a period of rebellion and turmoil (Arnett, 1999). Data show that self-esteem is usually low in this period and rises as people reach young adulthood (Bachman, O'Malley, & Johnston, 1978). Recklessness and sensation seeking also seem to decline after adolescence (Zuckerman, 1979).

These observable changes led to theories of child personality development, of which Freud's is historically the most influential. Psychologists trained in this tradition began to ask if the same kinds of developmental changes could be taking place in adulthood. If there were oral, anal, and phallic stages in childhood, might there not be later psychosexual stages for adults?

This investigation by analogy or extension is in the highest tradition of science. Physicists look for (and find) subatomic particles that in some respects parallel known particles. Cognitive psychology has benefited from computer models. Often this procedure can serve to refocus our perspective and allow us to "see" phenomena we have never noticed before but that are obvious once our attention is called to them. The discoveries of lifespan development may be equally convincing once they are made (indeed, we think they are).

The fact that aging and personality constitute a field in search of a phenomenon is itself an interesting phenomenon. What it seems to mean is that the changes in personality that occur in adulthood—if indeed there are any—are less dramatic than those of childhood. There are some stereotypes of old and young people, but these are notably inconsistent. Romantic idealism is thought to be characteristic of the young, but what about Don Quixote? Age purportedly brings a mellowing of the spirit—except to cranky old men.

We can agree that old people are less healthy than young ones and

that as they age many of them lose their hair, teeth, and hearing. We cannot seem to agree that they become more or less anxious, loving, or withdrawn. The field of aging and personality was intended to answer such questions.

One possible explanation for the lack of common knowledge about personality change in adulthood is that there *are* no changes. But before we jump to that conclusion, we must bear in mind the particular problems that a common-sense focus has in formulating ideas about aging. In order to detect a pattern, we need to see a phenomenon repeatedly. Some parts of what we see are the result of chance, some the result of an underlying regularity, and only repeated observations can tell which is which. The parents of a first baby are concerned with every change, not knowing what is normal development and what might be a sign of illness. By the third or fourth child, however, the pattern is familiar, and the experienced mother is something of an expert in child development.

But we do not live long enough to have repeated experience with adult development. We know our grandparents as old people, but we do not know what they were like as children. We can watch our parents grow old, but it is difficult to separate our own maturing perceptions from real changes in them. We can, of course, observe our own life and the lives of our friends, but we could expect to draw conclusions then only at the end of our lives. And we have grown up in a particular period of history whose twists and turns, rather than the aging process itself, may have made us what we are.

In short, common sense does not provide the distance, or perspective, from which to grasp any facts about adult development except the most obvious. Historical biographers, who can compare across the centuries, and anthropologists (Myerhoff & Simić, 1978), who can contrast cultures, can offer some insight. As psychologists, we base our approach on the premise that scientific measurement of personality characteristics in aging people can also provide a basis for answering this question. Quantitative investigations of large samples of people followed over a period of years can detect even very subtle changes with great objectivity. And psychologists have been concerned with issues of personality development long enough that we now have data following the same individuals over the greater part of their lives. This book is based, first and foremost, on the results of those studies. As we interpret them, they point clearly to the conclusion that personality forms part of the endur-

ing core of the individual, a basis on which adaptation is made to an ever-changing life.

It is only fair to warn the reader that the interpretations we make are not universally shared (although they are much more mainstream now than when we first proposed them). A number of theories of adult personality development have been proposed, and many researchers believe that they provide a more insightful account of adulthood than the stability view offers. The evidence on which the stability view rests, like all empirical evidence, is open to alternative interpretations, and a number of these have been advanced. In particular, as we will see in detail in Chapter 10, there is fundamental division on what exactly is meant by *personality*. It may turn out that the aspects we emphasize in personality are stable whereas what others emphasize changes (McAdams, 1994). Special attention will have to be paid to the issue of defining and measuring personality.

A Note on Psychotherapy

An account of some of our research published in the popular press was headlined, “Your personality—you’re stuck with it” (Hale, 1981). As clearly as anything else, that phrase illustrates how findings of stability are often seen and why they are frequently unpopular. Our findings seem to be read as a sentence of doom for all people who are unhappy with themselves. But the findings do not necessarily mean that at all. We hope to show in the remainder of this book that the process of aging in itself does not bring about substantial changes in personality and that most people change little from age 30 to age 80 in some of the most central aspects of their social and emotional makeup. Most people do not change—but that does not mean they *cannot* change. It does suggest, however, that the change will not come of itself, nor will it come easily. Effective psychotherapy or major life experiences (such as war or religious conversion) may profoundly alter us, but usually this occurs only if we are ready for a change and willing to work to make it happen.

What is “effective psychotherapy”? Each major school of therapy is able to claim its own victories (e.g., Chambless & Ollendick, 2001; Rogers & Dymond, 1954; Stuart, 1977) and research (VandenBos, 1986) confirms the value of psychotherapy in general. But virtually all psychotherapists would agree that real change in personality cannot be achieved at all without intensive and generally long-term efforts by

skilled professionals who create very special circumstances. We will return later to the issue of personality change associated with the use of therapeutic drugs like Prozac.

When Does Adulthood Begin?

The claim that personality is stable in adulthood needs some clarification. Who, after all, is an adult? Legal definitions of adulthood vary widely, not only by state but by function. A woman is legally of an age to marry at 14 in many states; the driver's license is withheld until age 16, voting until 18, and drinking until 21. Insurance companies, with actuarial wisdom, charge higher rates for drivers under 25. William James and a subsequent generation of hippies set age 30 as the dividing point between youth and settled adulthood.

Most psychologists probably consider college students—at least by the time of graduation—full-fledged adults; and in many respects, of course, they are. There is reason to think, however, that personality development continues, at least for some individuals, for several more years (K. M. White, Spiesman, & Costos, 1983). Studies that trace individuals from college age into later adulthood almost invariably report some changes in the average levels of personality traits and more fluctuations for individuals than are found in studies of individuals who are initially older (Finn, 1986; Helson & Moane, 1987; Jessor, 1983; Mortimer, Finch, & Kumka, 1981; McGue, Bacon, & Lykken, 1993). Haan, Millsap, and Hartka (1986) concluded from their study of personality that important changes may occur after high school:

Great shifts in personality organization are ordinarily thought to occur *during* adolescence, but these findings suggest that more marked shifts occur, not during adolescence, but at its end when most people make the profound role shifts entailed by entry into full-time work and marriage. (p. 225; emphasis in original)

That somewhat surprising conclusion is bolstered by recent longitudinal data showing that there are few changes in the mean levels of personality traits in the age range from 12 to 16 (Costa, Parker, & McCrae, 2000).

What is the nature of the changes seen in the decade of the 20s? When we compare the personality scores of college students to those of

adults on the NEO Personality Inventory (Costa & McCrae, 1985, 1989a), our measure of the five factors of personality, we find that students are somewhat higher than adults in Neuroticism, Extraversion, and Openness, and lower in Agreeableness and Conscientiousness. These differences suggest (although, as our discussion of cross-sectional research in Chapter 4 will explain, they do not *prove*) that college students mature and mellow a bit, becoming a little less emotional and flexible, but kinder and more responsible.

If we define adulthood as the period from age 18 on, studies like this make it clear that there is indeed adult development in several aspects of personality. We have adopted a different definition of adulthood. The data suggest to us that personality change is the exception rather than the rule after age 30; somewhere in the decade between 20 and 30, individuals attain a configuration of traits that will characterize them for years to come. From the perspective of the trait psychologist, adulthood begins at that point.

OTHER VIEWS: THEORIES OF CHANGE

A brief introduction to theories of change is an essential starting point for this book; much more will be said about the theories of Daniel Levinson and Roger Gould in Chapter 9. Readers may want to consult the original sources for more extended (and perhaps more objective) accounts of these theories.

Among major personality theorists, only two made significant contributions to lifespan theory: Jung and Erikson. The former left relatively unstructured ideas but was very influential in turning attention to the later years. Erikson, on the other hand, produced an elaborate and finely articulated theory that has become the basis of considerable empirical work (e.g., McAdams & de St. Aubin, 1998; Tesch & Cameron, 1987; Whitbourne, 1986a, 1986b). His theory of eight stages of life is probably the most important classic theory of adult personality development.

Jung, it will be recalled, objected philosophically to Freud's emphasis on sexuality. Jung proposed that sexual development was central to personality only in the young, in whom the function of procreation was vested. Past the age of 40 there must, he felt, be other sources of growth, other areas of concern to the individual. These were more likely to be

spiritual than sexual, and instead of revolving around the social function of procreation, they concerned the individual's relation to self. *Individuation* was the term he coined to reflect the continuing process of self-discovery and self-development, which he hypothesized would occur in the second half of life.

One of the key concepts in individuation was a balancing of opposing traits. Jung (1923/1971) conceived of the psyche as having two sets of opposed functions for dealing with internal and external reality: thought versus feeling, and sensation versus intuition. At any time in the life of the individual, one or another of these functions would be dominant and its opposite would be repressed. In order for there to be a full and complete expression of the self, however, the repressed side of the personality would also have to be allowed its chance. The psychological ideal was found in old age, when an integration of opposing functions would mark the culmination of development. Similarly, other personality structures, such as the persona (the mask one wears in social interactions) and shadow (rejected and unconscious aspects of the psyche) and the anima or animus (the feminine side of men or masculine side of women) must also be integrated. In general, this view of adult personality would lead to the expectation that the manifest characteristics of a young person should change markedly with age, either becoming their own opposites or moderating in degree as they were integrated with their complements.

Jung's theories were based on his experience with psychiatric patients and on his own experience of aging, and were buttressed by his scholarly studies of such obscure topics as alchemy and the *I Ching*. Few psychologists have claimed fully to understand his ideas, and few subscribe to them in toto, but some of the basic notions, such as continued and personalized development and the rise of repressed sides of self, have left a profound mark on subsequent theories.

Erik Erikson's (1950) views of adulthood, on the other hand, have become received wisdom in all their details. Erikson belongs to the group of ego psychologists (along with Erich Fromm, Karen Horney, David Rapaport, and Anna Freud) who began with classical psychoanalytic theory and in greater or lesser degree modified it to take into account features they believed had been slighted by Freud. Classical psychoanalytic theory is fundamentally biological in tone and leaves little room for the influences of culture or the individual's own efforts at growth and change. Erikson's solution to the problem of accounting for

social and environmental influences on personality was to argue that the traditional psychosexual stages that formed the backbone of Freudian personality development were paralleled by psychosocial stages. In addition to oral conflicts, the infant was faced with the resolution of the issue of basic trust versus mistrust, and each person's resolution of this conflict was heavily shaped by the cultural traditions that dictated methods of child rearing. Corresponding social issues were postulated for anal, phallic, latency, and genital stages.

Having made the transition from sexual to social stages, Erikson found that he was free to extend the social stages beyond the limit of sexual development. He hypothesized that the young adult needed not only genital gratification but also psychological intimacy in order to form the lasting bonds needed for the establishment of family life. He or she was required to resolve the crisis of intimacy or be left with a pervading sense of isolation. In the period of child rearing and adult careers, a new issue—generativity—became salient. Individuals who do not adopt an orientation that fosters growth in their children and community succumb to a sense of stagnation and meaninglessness. Finally, in old age the approach of death and the completion of life tasks leave the individual with the realization that his or her life is over and that there will be no second chances. Poorly resolved, this crisis leads to bitterness and despair; well resolved, it brings about a sense of ego integrity and an acceptance of both life and death.

Erikson's model of development is *epigenetic*, which means in part that the resolution of each crisis depends on the outcome of former resolutions. The best preparation for ego integrity is a life marked by intimacy and generativity (as well as desirable outcomes of the crises of childhood). But Erikson also admits a considerable element of changeability: At each crisis there is a possibility of new success or failure. The opportunity to redeem a misspent life at any age is one of the more inviting aspects of this theory.

Most of the empirical research on Erikson's stages has been confined to the period of adolescence and the transition to adulthood (e.g., Constantinople, 1969; Whitbourne & Waterman, 1979), but occasional studies have also been conducted on middle-aged and elderly people (McAdams & de St. Aubin, 1998). Erikson's ideas have been more widely adopted, sometimes in modified forms. Certainly the notion of stages of adulthood has become widespread, and the conception that orientation to life must change with point in the life course seems un-

questionable. (As Neugarten, 1968, notes, at some point in middle age, life begins to be measured in terms of time left to live rather than time lived.)

After the 1950s, research took the place of theorizing in this area. Even earlier one investigator had made significant contributions to the literature. Charlotte Bühler (1935) examined diaries and other personal records to chart the course of life in substantial samples of people. She noted a general "curve of life" including periods of growth, maintenance, and decline, and focused on the motivational changes that she believed occur with age. For the young person, instrumental strivings are central to daily activities; for the older person, these become much less important and are replaced by concerns for intrinsic values.

A large number of empirical studies were undertaken, in some cases to test the theories proposed by Erikson and Jung, in others simply to see what happened as people aged (Neugarten, 1977). Cross-sectional comparisons of young and old on a plethora of variables were conducted, retrospective accounts were collected from aged men and women, and a handful of longitudinal studies were launched. We will return to these studies in later chapters; they do not concern us here since they did not, in most cases, lead to theories of adult change. A few regularities were reported from the cross-sectional studies, including an increase in the level of introversion with age, but this change was rarely interpreted.

An exception to the atheoretical bent of these researches is found in the work of Neugarten and her colleagues. At least two significant theoretical concepts of change in adulthood emerged here, based in both cases on research using the Thematic Apperception Test (TAT), in which psychologists interpret the stories told by subjects in response to a standard set of pictures, some commonplace, some bizarre. Neugarten herself (1964) was responsible for the concept of *interiority*. She postulated that older individuals turn inward and consolidate their sense of self. The identity that the adolescent takes on and the adult acts on is further distilled in old age, and individuals become more and more like themselves. An increase in social introversion reported in some cross-sectional studies is sometimes taken as evidence of increased interiority, but even Neugarten (1968) admitted that there seems to be no regular change in social and emotional functioning with age. Some intrapsychic process not readily observable to an outsider must be meant by interiority.

Similarly, David Gutmann's (1964) concept of ego mastery styles is tied to intrapsychic changes inferred from TAT responses. Guttmann finds evidence in his TAT stories for three styles of mastery: active, passive, and magical. The first is seen in stories that show the hero taking forceful action to solve problems; the second is inferred when the hero accepts conditions as they are and adapts to them; the last, magical mastery, is seen in stories in which the hero distorts the situation or fails to see obvious dangers and problems. A "magical" solution has lost contact with reality.

Gutmann proposes that there is a universal developmental sequence in mastery styles. Young men use active mastery; middle-aged men, passive mastery; and old men, magical mastery. Originally found in a sample of men from Kansas City, the same pattern reappears, according to Gutmann (1970, 1974), in rural Mexicans and in the Highland Druze of Israel. A somewhat different pattern is found in women, a tendency to use passive mastery appearing in young women and active mastery in older ones. The crossover of the sexes is interpreted as change in masculinity-femininity, somewhat in conformity with Jung's notions of balancing. As women age they become more masculine in mastery style; as men age, they become more feminine.

It is imperative to note that an ego mastery style is the hypothesized basis for *experiencing* events; it is not the basis for overt action. Gutmann points out that among the Highland Druze the oldest men are the Elders, who have great power in decision making and are vigorous and decisive in running their community. The significance of a magical mastery style is thus far more subtle than it might at first seem. The discrepancy between the overt and observable and the unconscious, intrapsychic, or inferable has bedeviled personality psychology from the beginning; interiority and ego mastery styles are simply the gerontological version of far more pervasive perplexities.

George Vaillant (1977), a psychiatrist, proposed a theory of adult development based on the maturation of defenses. Defense mechanisms in one form or another have been central to many psychodynamic theories from Freud on; Vaillant built on this tradition by proposing that there are 18 basic defenses grouped into four levels that can be ranked in terms of psychological maturity. The least healthy are the psychotic mechanisms of Level I that alter reality for the user and include delusional projection, denial, and distortion. Level II mechanisms operate not by changing reality but by altering the user's distress, either experi-

enced or anticipated. These immature defenses include projection, schizoid fantasy, hypochondriasis, passive aggression, acting out, and dissociation. Neurotic defenses (Level III) include reaction formation, isolation, displacement, and regression. Mature defenses (Level IV), which integrate reality and private feelings, include humor, altruism, sublimation, suppression, and anticipation. Working with the Grant Study of Adult Development, Vaillant and his colleagues studied the use of these mechanisms in a longitudinal sample of 95 Harvard alumni. An extensive file had been maintained for each subject, including transcripts about how they had dealt with recent problems and stressors. Vaillant abstracted over 1,700 vignettes from these files, and raters, blind to the age and identity of the subject, evaluated each to determine the kind of defense mechanism used. (It should be noted that in the psychoanalytic tradition, virtually all behavior, even the most rational and mature, is ultimately defensive, so all vignettes could be scored.)

When the responses were divided into three periods (less than 20, 20–35, and over 35), there were systematic changes in the use of the different categories of defense mechanism. Between adolescence and early adulthood there were clear increases in the use of mature mechanisms and decreases in the use of immature mechanisms; neurotic mechanisms did not show differences. This trend continued, though more weakly, in comparing early adulthood (20–35) with later adulthood (over 35). The data led Vaillant to see adult development in terms of ever more mature forms of defense. Note that this theory is supported by data only for the early years of adulthood—indeed, it may provide further evidence that full adulthood is not reached until age 30. Whether maturation of defenses continues into later life is still a matter of conjecture.

In the 1970s two major theories of adult development were generated at about the same time. (Sheehy's *Passages* was based on early formulations of these.) Daniel Levinson and his colleagues (1978) wrote on the *Seasons of a Man's Life*, whereas Roger Gould (1978) called his work the study of *Transformations*. Both hold that there are distinct stages of personality development in young and middle adulthood, but the differences between the two theories are as notable as the similarities.

Levinson and his coworkers gave intensive interviews to 40 men—10 executives, 10 workers, 10 biologists, and 10 novelists (his work on women, published posthumously [D. J. Levinson & J. D. Levinson,

1996], claimed to find similar patterns of development). In cooperation with these subjects, the researchers wrote a biography for each man that focused on changes in what Levinson called the *life structure*. The life structure includes personality, but it also includes career, marriage, other relationships, values, and so on. According to the scheme worked out by these writers, adult life is divided into fixed and age-related stages. After the early career beginnings comes a period of reassessment at the age-30 transition. A much more searching reassessment comes at the midlife transition at age 40. In this period, often characterized as a midlife crisis, there is a period of inner turmoil that, according to Levinson, often seems to resemble neurosis. Dreams and aspirations that the young man had repressed in dealing with the realities of starting a career now come to the fore, and often a new career is launched that better addresses the long-denied needs.

Gould, whose version of adult development is more closely tied to personality, devised a version of psychoanalytic thought to explain the changes he thought he saw in adults. At the heart of the theory is the basic insecurity of the child, faced with the uncertainties and dangers of the world. To cope with these fears, the child adopts a set of beliefs that Gould called *illusions of safety*. Thus, children believe that their parents will always be there to take care of them; that there is no real evil or death in the world; that life is simple and controllable. Each of these illusions is comforting to the child, but each leads the adult to a distorted view of the world. Over the course of early adulthood, individuals must come to terms with these beliefs and abandon them to find a more realistic view of the world.

Gould considers that these illusions of safety, like psychodynamic defenses, are unconscious. Young adults are unaware that they assume that parents will always be there to help—in fact, they may explicitly deny such a belief. But, says Gould, they will act as if they harbored these delusions, and that is the key to the action of unconscious forces. Becoming an adult, attaining full adult consciousness, depends on outgrowing these illusions: a painful but salutary process. There is, according to Gould, a regular sequence in which the illusions are tackled and thus a rough correspondence to age, but there are fewer chronological absolutes in his system than in Levinson's.

Perhaps the most ambitious new theory of lifespan development is called SOC, for *selection, optimization, and compensation* (P. B. Baltes, 1997). This theory describes how individuals react to the opportunities

and limitations that life presents at each point in the lifespan. Selection refers to the choice of goals; optimization to their effective pursuit; and compensation to the use of alternatives when established methods fail. For example, an older person suffering a loss of vision may give up reading (selection) or may use higher-wattage lights (optimization), or may turn to audio books (compensation). Most people perceive a developmental curve (like Bühler's curve of life) in SOC processes: Our options and abilities increase from childhood into midlife, and decline thereafter.

SOC is not a theory of personality, but it is related to personality variables, because there are individual differences in the success with which SOC is applied. Some people manage their lives well; they have clear and focused goals, pursue them diligently, and show flexibility in their strategies. Other people seem aimless and feckless, unwilling to abandon strategies that have repeatedly proven ineffective. It is perhaps not surprising that measures of SOC are associated with Conscientiousness and Openness (Freund & Baltes, 1998). These associations lead to interesting questions about personality and the lifespan. For example, compensation is most important in the last phase of life, when health limitations are most common. Open people should fare better in old age, and the question of whether people become more or less open with age takes on added significance.

* * *

This review of the major theories of adulthood reveals their major strength and one of their weaknesses. The strength lies in the sheer attractiveness of the proposition that individuals continue to grow and change as adults. Surely all the experience of years counts for something! Surely the universal changes in health, appearance, and intellectual functioning have some parallel in personality! Surely there must be some hope for individuals whose current life is not the kind they would want to repeat for the next 50 years!

The first and in some respects the most troubling weakness with the theories is their mutual inconsistency. A number of very thoughtful and insightful observers of humankind have contemplated the course of adult life and have pointed out patterns they seem to see. But these patterns as often contradict as support one another. Gould sees his patients coming closer and closer to reality; Gutmann detects a retreat from real-

ity. Bühler sees the last phase of life as one of decline; Erikson sees it as the time for the development of wisdom. Levinson puts age limits on his stages of adult change; Neugarten hypothesizes a steady increase in interiority. The view that age brings maturation in forms of defense (Vaillant, 1977) conflicts with the conclusion that older people employ primitive defenses in dealing with stress (Pfeiffer, 1977). When observers from a dozen different perspectives see the same phenomenon, we begin to believe it is really there. When everyone reports something different, it is hard to know how to interpret it. And that is the time to turn away from personal impressions and look at the facts.

A Trait Approach to Personality

Conflicts among theories of adult development pale in comparison with conflicts among theories of personality. At least gerontologists can all agree that they are interested in what happens to people as the years pass. The only common denominator of personality psychology is the attempt to provide a psychological account of the person as a whole. But schools of psychology are so different in the phenomena they emphasize—from stimulus–response bonds to unconscious archetypes to cognitive maps—that the resulting theories of personality are often incommensurable. We cannot cover all these approaches fully, nor do we think they all deserve attention. Instead, we will concentrate on trait models of personality, both because they are compatible with a wide variety of theoretical approaches and because they have formed the basis for most research on personality. Personality measures, whatever their theoretical origins, are usually measures of traits.

Within the more limited scope of trait psychology there are still many controversies, but there is also evidence of real progress toward consensus in recent years, after a period of demoralization. Questions about the cross-situational consistency of behavior (Mischel, 1968), the validity of trait measures (Fiske, 1974), and even the objective reality of traits themselves (Shweder, 1975) had thrown the field into crisis in the 1970s; there was a widespread belief that personality psychology was on its last legs.

If we held the same skeptical view today, we would certainly not be

writing this book. The developments of the last 30 years have been little short of spectacular: Personality psychology has had a renaissance, with major advances in conceptualization, description, and measurement, and the Five-Factor Model (FFM) has brought order and understanding to the endless list of specific traits. Lifespan developmental psychology and longitudinal studies have made a major contribution to this rebirth by demonstrating the reality and significance of enduring dispositions.

The renaissance has been made possible only by a reformation in theory and research (Costa & McCrae, 1980c). Standards of research have become more rigorous, measures more critically scrutinized, replications demanded and provided. Correspondingly, theorists have become more willing to qualify and modify their positions on the basis of empirical findings. Perhaps for the first time, personality psychology has begun to be a real science (McCrae, 2002).

Clearly, an entire volume could be devoted to the state of personality psychology today. At best, this chapter can only give some sense of the issues and an outline of the arguments—and the data—we think make a case for a particular model of personality. The model we develop below is central to the remainder of this book, for when we say that personality is stable, we mean personality as defined by that trait model. We also discuss personality measurement later in this chapter, since measures and their interpretation allow us to check theories against the facts. We offer our preferred theories, models, and methods in this and the following chapter, but even in the current climate of emerging consensus the reader should realize that our preferences are not universally shared (J. Block, 2001). The issues we raise here are not put to rest—indeed, they recur throughout the remainder of the book. In Chapter 8, in particular, we contrast trait psychology with ego psychologies.

Perspectives on Human Nature

As every student of personality theory knows, three major schools of psychology have been reflected in theories of personality: psychoanalytic, behaviorist, and humanistic. Psychoanalytic theories of personality (Freud, 1933, 1900/1938) stress the individual's unconscious motivations, which must be inferred from such indirect sources as dreams, slips of the tongue, and fantasies. Behaviorist versions of personality theory (Dollard & Miller, 1950) limit themselves to observable behavior and invoke situational determinants, expectancies, and histories of rein-

forcement to explain behavior; more recently, social learning theories (e.g., Bandura, 1977) have acknowledged the role of cognitive processes in shaping behavior. Humanistic psychologies (Maddi & Costa, 1972), which arose in reaction to what were perceived as the irrational and mechanistic biases of psychoanalytic and behavioral theories, emphasize humankind's capacity to think, love, and grow. Each of these approaches has made valuable contributions to personality psychology, and we say more about them later, especially in Chapter 8. Our current concern, however, is to draw attention away from them to an alternative approach to personality that we feel deserves more credit than it is normally given.

The major schools have been interpreted as representing three different philosophies of human nature that, for the sake of contrast, are often depicted in oversimplified form. These caricatures of human nature may not do justice to the intricacies of the theories themselves, especially in their contemporary forms, but they do highlight the basic concerns of each perspective. From the psychoanalytic tradition, we might infer that the individual is basically irrational, driven by animal instincts, with rational control maintained only by the countervailing forces of socially induced guilt and anxiety. From a behaviorist perspective, the individual is seen as less ominous and unpredictable; indeed, human nature is largely or wholly the result of experiences in the social environment that shape and reward certain behaviors. People are reactive, habit-bound creatures of the environment they live in. Humanistic psychologists endow people with a far more pleasing aspect; love, creativity, and play are thought to be the quintessential features of human nature, and both irrationality and rigidity are interpreted as signs of the destructive influence of society.

How is it that theorists who look at the same phenomenon—human nature—draw such different conclusions about it?

One possibility that may account for this selective attention to the facts about human nature is that people are different. Freudians hold that the human species is naturally aggressive, and they have no trouble pointing to infamous examples of the kind of humanity they envision. Rogerians (Rogers, 1961) believe in the capacity for openness and love, and they, too, find exemplars of their ideas. (Abraham Maslow, 1954, studied such self-actualizing individuals as Abraham Lincoln and Eleanor Roosevelt as a kind of counterbalance to the case studies of psychiatric patients that had formed the basis of Freudian psychology.)

Despite the claim that people are able to transcend their environments, it is all too easy to find individuals who seem to be wholly a product of a history of reinforcements.

For decades the debate has raged about which of these is the true image of human nature. The idea that they might all contain an element of truth is a truism that has captured the imagination of almost no one. But it forms the basis of another school: the psychology of traits, or individual differences. This perspective resists facile and sweeping generalizations:

Are people basically selfish? Some are, some aren't.

Are human beings intrinsically creative? Some are, some aren't.

This position, which emphasizes the consistent differences of individuals, has always played a major role in both common sense and academic psychology. It might also be seen as a philosophy of human nature itself, one well suited to a pluralistic society (Costa & McCrae, 2000c). But it has not received much attention as a philosophical position, and from the beginning trait psychology has been regarded as a relatively minor part of personality theory: not as a fourth school, but only a set of personality measures, a few isolated studies, an appendix to one of the "true" schools.

One reason for this is that it has been possible to view individual differences simply as one aspect of an all-encompassing theory. To the degree that each of the other schools attempts to account for at least some individual differences, each incorporates the trait model. Classical psychoanalytic theory (Freud, 1900/1938), for example, proposed that the resolution of psychosexual conflicts in childhood and the development of characteristic defenses led to enduring character traits, such as those typifying the anal personality. For the psychoanalyst, neatness, punctuality, thrift, and cleanliness are viewed as the outcome of a fixation during the anal-retentive stage of development. Maslow, a representative humanist, defined individual differences in terms of levels of basic motivation. Social learning theorists (Bandura, 1977), the contemporary descendants of behaviorists, might explain characteristics like masculinity or femininity as the result of role-modeling and socialization processes.

In short, every theory of personality is concerned with the differences as well as the similarities among people. But the concern tends to

be secondary. Maddi (1980), who reviewed the major types of personality theory, discussed individual differences in terms of what he calls peripheral characteristics, in contrast to the core characteristics that form the heart of personality theory. Trait psychology would seem to be all periphery and no core; why settle for that when other schools promise both?

The answer is, in part, because other schools are less than fully successful in delivering a good model of individual differences. Each school leads to an emphasis on certain characteristics, often to the exclusion of others. Freudian theory, for example, is primarily concerned with neurosis and impulse control. It has an elaborate system for describing the varieties and degrees of maladjustment. But it fails to say much of significance on other individual differences, such as why some individuals are introverted, some extraverted. Psychoanalysts might well claim that, from their point of view, differences on that dimension are trivial. Other traits are trivial from other points of view. Only a system that grants the first place to individual differences will attempt to provide a comprehensive list of traits. And comprehensiveness is essential if we are interested in the area of personality and aging. How can we specify what characteristics do and do not change unless we know the full range of characteristics to look at?

Historically, trait models appear to be by-products of personality theorizing. Wouldn't it be possible to begin with traits and build up a theory of the whole personality system? That would provide the dynamic context in which traits operate and should satisfy the need many psychologists feel for a more process-oriented approach to personality. In recent years we have tried to do just that, by proposing a Five-Factor Theory of personality (FFT; McCrae & Costa, 1996, 1999). That theory is based on trait research, and a description of FFT is best postponed until the findings have been presented. In the meantime, we can proceed with a purely phenotypic approach to traits.

BASIC PRINCIPLES OF TRAIT PSYCHOLOGY

In common speech, when asked to describe someone, we generally rely on trait terms. We say that one person is hardworking and aggressive, another is hostile and unimaginative, a third is shy but loyal to friends. Learning how to use such words properly, as applied both to others and

to oneself, is an important part of language acquisition. As naïve psychologists, all English-speaking people subscribe to a trait theory of psychology. Personality trait terms are also central to most or all other human languages (Goldberg, 1981), probably because the language of traits is so useful in facilitating interaction with other people that every culture has invented it.

As scientists, trait psychologists have tried to go beyond the common sense view of traits. But the best of them, like Gordon Allport (1937), have also acknowledged the contributions of the natural language of traits. Human beings have been trying to understand and describe each other for thousands of years, and it would be presumptuous indeed to suppose that psychologists could start from scratch and come up with a better system in a few years. Trait researchers have borrowed from the common-sense system in two respects: First, trait conceptions of personality have often begun by attempting to spell out the assumptions behind the use of trait words. Second, trait measures have been based in varying degrees of immediacy on the language of traits that has been built up over centuries. We return to that aspect later in this chapter.

We can define traits as *dimensions of individual differences in tendencies to show consistent patterns of thoughts, feelings, and actions*. This is a phenotypic definition; in essence, it tells us what traits look like and how we can recognize one when we see it. A different, deeper definition is also possible, but that will have to be postponed until FFT is introduced. We can go a long way with the present definition. Characterizing traits such as shyness and trust as *dimensions of individual differences* means that people can be ranked or ordered by the degree to which they show these traits. Some people are very trusting, most are moderately trusting, but a few are quite suspicious. In fact, all the traits that we are concerned with in this book are found in varying degrees in all people, with distributions that approximate the familiar normal curve. Although some personality theories describe types—distinct groups of people characterized by a unique configuration of features—current research supports the view that most so-called types are merely extreme scorers on continuously distributed trait dimensions (McCrae & Costa, 1989b; Widiger & Frances, 1985).

The more of a trait people have, the more likely they are to show the behavior it disposes toward, and thus the more frequently we are likely to see it. Similarly, the more the trait characterizes them, the more

intensely they act and react in relevant situations. A very gregarious individual really likes to be around people, and frequently is. Frequency and intensity of the appropriate acts and feelings are the major signs from which we infer the level of a trait.

The word *tendencies* in our definition emphasizes the fact that traits are only dispositions, not absolute determinants. A great number of other factors go into the choice of a particular action or the occurrence of a particular experience. Gregarious people like to talk, but they normally do not chatter on during moments of silent prayer. Well-adjusted people may not worry much, but even they are likely to be anxious when awaiting the results of a job interview or a medical procedure. The requirements of the social roles we play, the facts of the current situation, the mood of the moment, and acquired habits all join in shaping the choice of a particular act, word, or emotional reaction.

This is a crucial point, because it was once considered a reason to discard the trait approach altogether. In 1968, Mischel published an enormously influential summary of the literature in which he argued that personality traits account for only about 5–10% of individual differences in actual behavior in any specific instance. This disappointingly small contribution seemed to suggest that the whole enterprise of trait theory and assessment was a fool's errand.

But, in fact, that finding is just what a thoughtful trait theorist would expect. Suppose we use the trait of warmth to predict whether an individual will smile at a stranger (Costa & McCrae, 1998b). We know that traits operate on a probabilistic basis and that people vary from low to high, with most in the middle of a bell curve. So we might predict that only 10% of the coldest research subjects would smile when we bring a stranger into the laboratory whereas 90% of the warmest subjects would; others would fall somewhere in between, perhaps as shown in Table 2. If these are the numbers we actually observed in the laboratory, we would be overjoyed at how well trait theory had predicted behavior. And yet, in this instance, the correlation between warmth and smiling is only .41, meaning that only about 16% of smiling behavior can be predicted from knowing the subject's trait level.

If traits are such weak predictors of behaviors even when they work perfectly, what is their attraction? As Epstein (1979) famously showed, the answer is that we are rarely interested in predicting a single behavior. We usually want to know how people act in general, and if we average behaviors across many instances, we can predict the average with

TABLE 2. Hypothetical Distribution of Smiling Behavior at Different Levels of Warmth

Level of warmth	Total subjects	Smiles at stranger?		Percentage smiling
		Yes	No	
Very high	70	63	7	90
High	240	168	72	70
Average	380	190	190	50
Low	240	72	168	30
Very low	70	7	63	10

Note. Adapted from Costa and McCrae (1998b).

greater and greater precision. Table 2 tells us that on any given occasion, whether a person average in warmth will smile or not is a coin toss. But it also tells us that we know with great precision that, in the long run, that person will smile about half the time. And since, as we will see, traits endure for long periods of time, their ability to predict long-term patterns of behavior is extraordinary.

In the short run, though, it is helpful to know how useful traits are in the real world, not the hypothetical distribution in Table 2. Fleenor (2001) has recently reported some interesting data relevant to that question. He asked a group of college students to record on a Palm Pilot how they had been behaving in the past hour—how talkative, hard-working, or cooperative they had been. They repeated this process five times a day for 2 weeks. He had two major findings: First, there was about as much variation within each person as there was between people—each person tended to show a wide distribution of talkativeness, hard work, and cooperation from time to time. Second, if half the ratings were averaged, they predicted the average of the other half almost perfectly. In other words, each individual had a characteristic mean level, but all people varied around that mean under different circumstances. Summer is always hot and winter cold, but there is a wide range of temperatures within each season.

By *patterns of thoughts, feelings, and actions*—to return to our definition of traits—we mean to indicate both the breadth and the generality of traits. Traits should be distinguished from mere habits: repetitive, mechanical behaviors such as smoking or driving fast or saying “You know” after every sentence. Habits are specific learned behaviors; traits

are generalized dispositions, finding expression in a variety of specific acts. Habits are thoughtless repetitions of earlier behavior; traits often lead people to develop entirely new behaviors, sometimes after much thought and planning. Driving fast may simply be a habit, perhaps learned from observing the way friends or parents drive. But if the fast driver also likes loud music and roller coasters, and perhaps experiments a bit with drugs, we begin to see a general pattern we can identify as excitement seeking (Zuckerman, 1979). An excitement seeker may spend weeks planning for a trip to Las Vegas, say, or she may decide to go on the spur of the moment. But in either case going to Las Vegas is not likely to be simple habit. In many respects traits resemble motives rather than habits, and it is often unclear whether a disposition such as excitement seeking should be called a trait or a motive. *Trait* appears to be the broader term, indicating motivational, stylistic, and other aspects of human consistency.

The *consistent patterns* that indicate traits must be seen over time as well as across situations. This means that traits are to be distinguished from passing moods, transient states of mind, or the effects of temporary stress and strain. If an individual is anxious and hostile today but calm and good natured tomorrow, we attribute these emotions to the situation—perhaps pressures at work or a quarrel with a spouse. Only when emotion, attitude, or style persists despite changes in circumstances do we infer the operation of a trait. Measures of personality traits are expected to show high retest reliability when administered on separate occasions days or weeks apart, because traits are characteristic not of situations, seasons, or times of day but of the individual at a particular point in his or her life.

Short-term consistency does not preclude long-term change. In fact, there could be neither stability nor change without a consistent pattern to begin with. From a trait perspective, personality change means that one consistent pattern is replaced by another. We review the evidence for such changes in Chapters 4–7.

Note that our definition of traits says nothing about their origins. Psychodynamic theories emphasize interactions of children with their parents; social psychological theories may point to peer pressures. There is increasing evidence for the importance of genetic influences on many traits (Plomin & Daniels, 1987; Jang, Livesley, & Vernon, 1996). Whatever their origin, by adulthood individuals can be characterized in terms of a variety of traits, and research on personality and aging can

proceed from that point. We will, of course, need to return to this issue when we formulate a trait theory of personality.

Similarly, we need not understand the physiological basis of traits in order to chart their development in later life. Most contemporary psychologists would acknowledge that there must be some neurophysiological or hormonal basis for personality, and some specific mechanisms have been proposed (e.g., Cloninger, Svrakic, & Przybeck, 1993). But to date the evidence in favor of these theories is limited, and it is unlikely that we will ever find a specific region of the brain that controls Neuroticism, or a neurotransmitter that accounts for Extraversion, or a single gene for Openness. The trait names do not refer to the underlying physiology but to the abstract consistencies in the ways people act and experience and to whatever complex psychological causes they may have.

How Many Traits? Which Ones?

Any number of human characteristics fit the definition of traits. There are physical traits (*tall, healthy*), ability traits (*intelligent, musical*), and social traits (*rich, famous*). Our interest is in personality traits, which describe emotional, interpersonal, experiential, attitudinal, and motivational styles. One way to get a sense of the scope and variety of personality traits is by consulting the dictionary: The English language provides an embarrassment of riches to the trait psychologist. So important are descriptions of people that thousands of terms have evolved, distinguishing minute shades of meaning that delight the poet and confound the empirical scientist. One major approach to personality has focused on analyses of natural language trait terms (John, Angleitner, & Ostendorf, 1988).

The more common alternative has been to begin with psychological theory and the insights of trained clinicians and researchers. After all, we do not expect chemists or anatomists to base their taxonomies on lay terminology, and many psychologists have disdained the common vocabulary of traits.

Theories of particular characteristics have often lead to the development of trait measures; hundreds of scales have been created in this way. Rokeach (1960) developed a questionnaire to measure his concept of dogmatism, and Crandall (1981) proposed a trait preference list to assess Alfred Adler's (1938/1964) concept of social interest. There are

dozens of scales intended to measure depression and others to assess closely related concepts such as loneliness and hopelessness. Many of these scales are technically excellent instruments, but their sheer number makes it difficult to comprehend the scope of personality variables.

In an attempt to bring some order to the field, several researchers have proposed trait systems, sometimes based on elaborate theories, sometimes simply reflecting a judicious choice of traits. Most have developed personality inventories to measure simultaneously all the traits in the system and thus provide a fuller personality profile than any single scale could give.

Jung's (1923/1971) theory of psychological types became the basis of many instruments. In particular, scales to measure the two Jungian attitudes of Introversion and Extraversion quickly became popular. In one of the earliest applications of factor analysis to personality research, J. P. Guilford and R. B. Guilford (1934) analyzed measures of Introversion–Extraversion and found that several different traits were represented. J. P. Guilford continued his factor analytic studies of personality and eventually developed an inventory of 10 traits, the Guilford–Zimmerman Temperament Survey, or GZTS (J. S. Guilford, W. S. Zimmerman, & J. P. Guilford, 1976). Another factor analyst, Hans Eysenck, noted that Extraversion was one of two fundamental dimensions that recurred in analyses of many personality inventories. The second dimension was emotional instability or maladjustment; because it was seen most clearly in individuals traditionally diagnosed as neurotics, he called this dimension Neuroticism. Later H. J. Eysenck and S. B. G. Eysenck (1975) added a third dimension, Psychoticism, to their list of basic personality dimensions. In addition to Introversion–Extraversion, Jungian psychology calls attention to two other contrasts: Sensation versus Intuition, and Thinking versus Feeling. The Myers–Briggs Type Indicator, or MBTI (Myers & McCaulley, 1985), measures these, as well as a contrast between Judging and Perceiving.

Henry Murray, whose 1938 volume *Explorations in Personality* is one of the classics of the field, was a psychodynamic theorist who viewed motivation as the key to personality. He and his colleagues at the Harvard Psychological Clinic identified a list of needs or motives they considered necessary for a reasonably complete description of the individual's personality. Needs for achievement, affiliation, and nurturance are among the list of 20 or so needs. Many instruments have been created to measure these needs, of which perhaps the best is Douglas Jack-

son's (1984) Personality Research Form, or PRF. Jack Block (1961) led another team of dynamically oriented psychologists and psychiatrists in an effort to develop a standard language for personality description. Instead of adopting a particular theory, their items were intended to provide a theoretically neutral language of description, useful to clinicians or researchers of any school. Further, they made a deliberate and sustained attempt to be comprehensive—to include all important personality characteristics. This is a crucial feature to which we return later.

The psychiatrist Harry Stack Sullivan formulated an interpersonal theory of psychiatry that has had influences on two important trait systems. Students of interpersonal behavior (Leary, 1957; Wiggins, 1979) discovered that most traits that described styles of interacting with others could be arranged in a circular order around the two axes of Love, or affiliation, and Status, or dominance. Figure 1 presents a diagram of this Interpersonal Circumplex. Individuals high on one trait (e.g., gregariousness) tend to be high on adjacent traits (dominance and warmth) and low on traits on the opposite side of the circle (aloofness). Con-

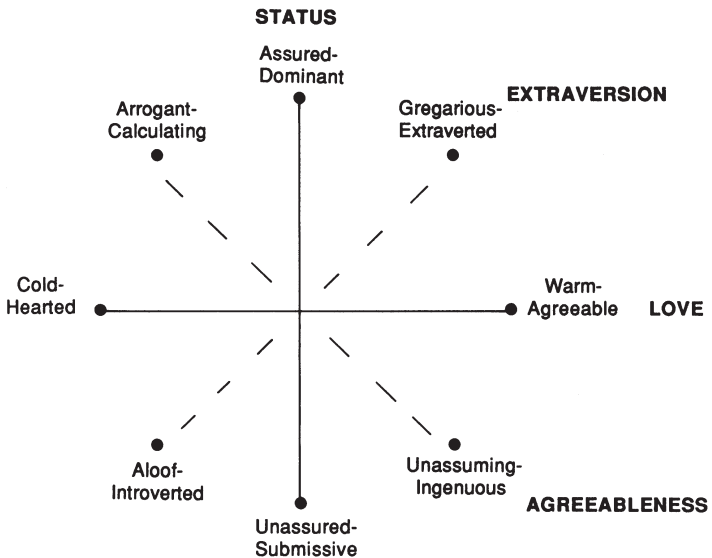


FIGURE 1. An adaptation of Wiggins's version of the Interpersonal Circumplex. The solid lines indicate the dimensions of Love and Status used in Wiggins's system; the broken lines indicate the dimensions of Extraversion and Agreeableness. Adapted by permission.

cepts of interpersonal interaction were also central to formulations of personality disorders (Millon, 1981; Widiger & Frances, 1985) and were ultimately incorporated in the 1994 edition of the American Psychiatric Association's official diagnostic system, the DSM-IV. These personality disorders can be interpreted as maladaptive variants of normal personality traits.

The Quest for a Unified System

Historically, competition among these systems has been a preoccupation of personality psychologists, and the lack of a single, unified model has significantly hindered progress in the field. Researchers from different schools could not easily communicate, so results could not be compared. A study of age changes on the GZTS would not speak to the concerns of adherents to Murray's system of needs (discussed above), so a separate study of age changes in needs might be required. Research efforts were at best inefficient.

If all the different systems really identified distinct parts of personality, as they appeared to do in principle, the search for a comprehensive yet parsimonious model would be futile. But numerous studies have shown that there are important overlaps among the various systems. Although needs may seem to be quite different from personality disorders and from Jungian functions, the measures of these concepts turn out to have much in common. Neuroticism and Extraversion play a large role in many personality inventories—so much so that Wiggins (1968) called these the “Big Two.” Unfortunately, the correspondences are often obscured by the labels psychologists have chosen for their scales. Two measures with the same label may measure different traits, and two measures of the same trait may have very different labels.

The identification of new traits might seem to be the equivalent of the entomologist's classification of insects or the astronomer's cataloging of stars. In fact, it is something quite different. It is easy enough to identify a beetle from its picture and description or to characterize a star by its exact position in the sky. But we cannot pin up specimens of anxiety or personal insecurity or Neuroticism or guilt-proneness and compare them, nor do we have a map of the “personality sphere” that we can use to check the location of a trait. How do we know whether the traits in one system are the same as those in another system under different names?

The only way we can tell if a scale measures a totally new trait or is merely some variation on a familiar one is by measuring a group of people on the proposed new trait as well as on all the older, known traits and seeing whether the measures of the new trait really give us some new information. (The technical name for this process is *discriminant validation*.) Since there are hundreds of scales in use, it is virtually impossible to measure all of them on an adequate sample of people and so settle the issue once and for all. But if we piece together all the research of the past 50 years in which a few traits have been measured jointly and compared, we can begin to see the outlines of the forest we are in. And once we have some notion of the major groups of personality traits, it becomes much easier to judge the contribution of new candidates. In short, we will have begun to map out the territory of personality and obtain the equivalent of the astronomer's map of the heavens and the entomologist's taxonomy of insects.

The major tool in this endeavor has been factor analysis, a mathematical technique for summarizing the associations among a group of variables in terms of a few underlying dimensions. The degree to which two traits go together, or covary, is assessed by the correlation coefficient, which is fairly easy to interpret. But as the number of traits increases, the combinations soon surpass the ability of the unaided mind to comprehend: When 35 trait scales are examined there are 595 different correlations; 50 scales yield 1,225 correlations. Factor analysis offers a way to determine how to group together sets of traits that are all related to each other and unrelated to other sets of traits. For example, anxiety, anger, and depression cluster together as part of the broader domain of Neuroticism; they are all relatively independent of traits such as sociability and cheerfulness, which are part of Extraversion.

It is its ability to condense and organize great quantities of information that makes factor analysis a natural instrument for researchers interested in studying the organization and structure of traits. As an analytic tool, however, factor analysis is not infallible. Different researchers can arrive at different conclusions from the same data when they use factor analytic techniques differently. Our view is that factor analysis should be guided by what we already know about personality structure, just as experiments are guided by theoretically derived hypotheses. Used in this way, factor analysis has led to solid and replicable findings about the structure of traits.

Natural Language and the Five-Factor Model

One way to avoid the controversies of competing psychological theories is to turn to the description of personality offered by natural languages such as English. This approach draws on the folk wisdom embedded in the language, and the resulting system will be readily understood by all speakers of the language: You must be trained in Freudian theory to know what an anal-retentive character is, but everyone understands what *stubborn* and *stingy* mean. Most importantly, however, the argument can be made that natural language can be used to define the full scope of individual differences in personality. Personality differences are important in getting along with others, in work and play, in maintaining traditions and creating new ones. Every culture must have evolved words to represent these differences, and over the course of centuries every important attribute or trait will surely have been noted and named (Norman, 1963). In their study of an unabridged dictionary, Allport and Odbert (1936) found about 18,000 such trait-descriptive terms—more than enough to occupy personality psychologists.

In fact, far more than enough—too many. Some screening, some clustering of synonyms, is needed to make this language manageable. Allport and Odbert began that job by identifying about 4,000 terms that most clearly referred to personality traits. The next step was taken by Raymond Cattell (1946), who formed groups of synonyms from the 4,000 words, eventually condensing them into 35 clusters. These were used in a study of personality ratings; the 35 scales were factored and 12 dimensions were identified. Combined with four additional dimensions Cattell had found in studies of questionnaires, these became the basis of Cattell's personality inventory, the Sixteen Personality Factor Questionnaire, or 16PF (Cattell, Eber, & Tatsuoka, 1970).

Our contributions to research on personality structure began with analyses of 16PF data on a sample of volunteers participating in a longitudinal study of aging conducted by the Veterans Administration in Boston (Costa & McCrae, 1976). After considering a range of possible solutions, we settled on what seemed to be both simple and adequate: a three-factor model. Two of the three dimensions were so similar to Eysenck's Neuroticism and Extraversion that we adopted those names; the third we called Openness to Experience.

Rokeach (1960), who was interested in the structure of ideas, had written about dogmatic individuals who were not open to new ideas or

values. The clinician Carl Rogers (1961) had observed that some of his clients were out of touch with their own feelings, and he proposed that openness to inner experience was a criterion of good mental health. Coan (1972) created a measure of openness to fantasy and aesthetic experience, and Tellegen and Atkinson (1974) noted that the capacity to become deeply involved in experiences, which they called *openness to absorbing experience*, was characteristic of individuals who were easily hypnotized. We saw all of these traits as part of a single domain of Openness.

We created an inventory to measure specific traits, or facets, of each of these three domains and conducted research with it for several years. Although it worked well in many respects, it shared with most other trait systems the limitation that it was not comprehensive. We knew that some traits, such as self-discipline, were not represented, but like other researchers we had no clear notion of what other traits might also be lacking. Ours was a good model as far as it went—but how far was that?

The quest for comprehensiveness led researchers back to the natural language. As long ago as 1961, two U.S. Air Force researchers, Ernest Tupes and Raymond Christal, had conducted a series of studies using the 35 rating scales devised by Cattell (Tupes & Christal, 1961/1992). Instead of the 12 factors Cattell had reported, they found only five—but they found them in several different samples. Warren Norman (1963) recognized the significance of this finding and, after replicating the five factors in his own studies, suggested that they constituted an “adequate taxonomy of personality traits.” Interest in this model waned, however, until Goldberg (1981, 1982) started the project from scratch, going back to the dictionary and forming his own sets of synonyms—and replicating the Five-Factor Model (FFM) once again.

Norman had called the factors Extraversion or Surgency, Agreeableness, Conscientiousness, Emotional Stability, and Culture. When, in the 1980s, we began to reexamine this system, we found striking parallels to our own. Norman’s Surgency was clearly our Extraversion, and his Emotional Stability was simply the opposite pole of our Neuroticism. If Norman’s Culture matched our Openness, then our three-factor model could be nested within the FFM and the claim that the latter was truly comprehensive would be considerably strengthened. Using adult men and women from the Baltimore Longitudinal Study of Aging (BLSA), we confirmed these hypotheses (McCrae & Costa, 1985b,

1987) and quickly became advocates of the FFM. We developed scales to measure Agreeableness and Conscientiousness to supplement our original instrument and published it as the NEO Personality Inventory, or NEO-PI (Costa & McCrae, 1985, 1989a).

* * *

If the FFM is a truly comprehensive taxonomy of personality traits, it has profound implications for the study of aging and personality. The task of looking for age changes, for example, is immensely simplified. Instead of being obliged to consider separately each of the 18,000 traits named in the English language and the hundreds of scales created by psychologists, we can concentrate on five broad groups of traits. By sampling traits from each domain, we can get a reasonable idea of how all traits are related to age.

To the extent that existing measures can be classified in terms of a taxonomy, the FFM can be used to summarize the existing literature and to look for similar patterns across different instruments. Further, the comprehensiveness of the model would ensure that our search for age changes was exhaustive. We would not need to worry that we had inadvertently overlooked some trait that shows striking age changes. If we see maturation in some traits in a domain, we may wish to look carefully for changes in other traits in the same domain. Studies of aging often require decades of data collection, so researchers need to be very selective in the variables they choose to study. The FFM can help us formulate plausible hypotheses about which traits must be represented and which can be safely ignored.

But the FFM has many other applications in the psychology of aging because it provides a framework for systematic research. Whether we are interested in the so-called midlife crisis, preferred age of retirement, adjustment to widowhood, or risk factors for dementia, we know we have fully considered the possible role of personality if and only if we have measured all five factors.

But is the claim to comprehensiveness justified? Does the FFM in fact encompass all the competing systems? This is an empirical question, and we must postpone our answer until we have considered how personality traits should be measured.

Measuring Personality

FROM CONCEPTS TO DATA

The whole trick of science is to test ideas against reality, and in order to do that, something has to be measured. We may suspect that salespeople are likely to be extraverts, but in order to prove it we have to be able to measure Extraversion and to convince other people that we are measuring it successfully. If we can come that far, it is a relatively easy matter to see whether salespeople score higher than others on our measure, although we also have to be concerned with the representativeness of our sample, the magnitude of the difference we find, and any number of other problems in interpretation.

In many respects, trait theories of personality are among the easiest to operationalize (as the process of finding suitable measures for theoretical concepts is called). Traits are generalized dispositions, so we know we must look for evidence of consistent patterns of behavior or reactions across a range of situations. Traits endure over time, so we know that our measures must give us about the same results when applied at two different times, say, a few weeks apart. Individuals vary in the degree to which they can be characterized by a trait, so we know our measure must show a range or distribution of scores.

Based on what we know about the domain, we would be able to measure Extraversion by getting answers to the following questions:

- How many close friends does this person have?
- How much does this person enjoy parties?

- Is this person often the leader of a group?
- How active and energetic is this person?
- How much does this person crave excitement?
- Is this person usually cheerful?

If we come up with a reasonable system for answering all these questions, say, on a scale from 1 to 5, we can add up the numbers and assign a total score for Extraversion. Once we have the answers for a sample of people, we could begin to check the utility of our measure by asking some simple questions. First, do these items generally agree in the picture they give of the individual? That is, do people with many close friends also tend to be energetic and cheerful? This is called *internal consistency*, and it provides evidence that the different behaviors or reactions are actually part of a pattern, not just a collection of unrelated activities. If the items lack internal consistency, either we have a bad set of questions or we have misunderstood the nature of Extraversion. Second, do the scores people get on one occasion parallel the scores they get at a later time? Are the extraverts of January still extraverts in March? This is called *retest reliability* and demonstrates that what is being measured is a trait, not a temporary state or mood of the individual.

There is little dispute about any of this. Most psychologists agree on what kinds of answers we need; the problem comes in deciding how we should get the answers. How do we know how much a person enjoys parties? We could ask the person for a self-report, but he or she may never have thought about it, or may lie to us, or may unconsciously distort the answer. We could ask friends for peer ratings, but they may not know how the person really feels and may have seen only one side of his or her personality. We could watch the person at a party and note how much he or she smiles, how long he or she stays, how excited and enthusiastic he or she seems to be—but his or her behavior at this party may be atypical. Perhaps the person has a toothache or does not get along with the host.

Behaviorally oriented psychologists would prefer the third strategy, behavioral observation, since it seems to be the most objective and scientific. But, as we have pointed out, the contributions of dispositions to observable behavior are often overshadowed by the press of circumstances, and studies have shown that behavior has to be observed over many occasions before reliable inferences can be made about traits (Epstein, 1979; Fleeson, 2001). In principle, repeated observation is per-

fectly reasonable; in practice, it is prohibitively expensive and time consuming.

Furthermore, observations of behavior, regardless of their number, may not be the most accurate source of answers. If we are really interested in an individual's dispositions, how the person feels may be much more relevant than how he or she acts. It may be necessary for him or her to put on a front, to play a role, to *seem* instead of *being*. A simple question asked in a way that allows for a candid answer may give a far truer picture of real dispositions. In some areas, such as happiness, it can reasonably be argued that personal opinion is the only meaningful criterion.

It would hardly be an exaggeration to say that trait psychology floundered over these issues for most of its history. The great majority of researchers have been content to hand out questionnaires and trust the self-reports of their respondents, but their conclusions have been dismissed time and again because of objections about the kinds of errors to which self-report instruments are prone (Berg, 1959; Edwards, 1957; Fisher, Schoenfeldt, & Shaw, 1993). Behaviorists have scoffed at the idea that paper-and-pencil measures can be a substitute for the observations of real behavior. Psychodynamic theorists have objected that unconscious defense mechanisms or more or less conscious tendencies to present a socially desirable image of oneself make self-reports suspect. Social psychologists have contended that what counts is the perception of behavior by significant others and that, in consequence, only ratings should be trusted.

In the past 20 years these controversies have been resolved, at least to the satisfaction of most psychologists. We know only too well that any single method of measurement, any source of answers to our questions, is subject to certain kinds of error and that consequently the surest results are those that can be seen using several different methods, as Campbell and Fiske (1959) noted long ago. But we are also beginning to see convincing evidence that, properly employed, many methods can give a reasonable approximation to the truth. A growing body of studies has shown that different methods do lead to the same conclusions about people.

One line of research has been concerned with correspondence between objective, observable behavior and self-reports. In one study (McGowan & Gormly, 1976), researchers followed a small group of college students and clocked their speed of walking. The fastest walkers

were the ones who reported the highest activity level and who were rated by their peers as most energetic—not a particularly exciting result for the average reader, but a revelation for those psychologists who thought that self-reports were totally untrustworthy. Similarly, Small, Zeldin, and Savin-Williams (1983) recorded altruistic and dominant behaviors of small groups of teenagers at a summer camp and also asked the campers to rate each other on these two dimensions. They found remarkably high agreement between their behavior counts and the peer ratings. Another version of this line of research (Epstein, 1977) has demonstrated that as the sample of behavior increases, the correspondence between observations and self-reports increases to quite respectable levels. Research of this kind continues to support the validity of self-report measures (Funder & Sneed, 1993), with an ever-increasing understanding of the reasons why self-reports and behavior ratings do not correspond perfectly (Moskowitz, 1988; Wright & Mischel, 1988).

Self-Reports and Observer Ratings

The direct observation of behavior in the laboratory or in natural settings is invaluable as a way of anchoring our scales to rigorous scientific measurements, but as a technique of personality assessment it is both cumbersome and expensive. It would be virtually impossible to make sufficient observations on sufficient subjects over a sufficient number of years to approach the study of age and personality in this way.

Fortunately, there are alternatives. Self-report is the most widely used method; here individuals are asked to describe themselves by answering the items of a questionnaire or checking off adjectives that apply to them. Individuals have spent a lifetime getting to know themselves, and they alone can draw on the subjective experience that tells them whether they truly enjoy an activity or, deep down, harbor vague doubts or anxieties. Provided that respondents are honest with themselves and with the researchers, self-report is probably the best way to measure personality.

That proviso, however, has always been disputed: Subjects may deceive themselves, or they may simply lie. For researchers concerned about these possibilities, there are three options honored by tradition: projective methods, which are supposed to allow the investigator to see behind the individual's facade; validity scales designed to detect faking or lying and adjust self-reports accordingly; and ratings from knowl-

edgeable informants. We postpone until Chapter 8 an evaluation of projective methods and focus for the moment on the other two approaches.

Observer Ratings

Ratings can be obtained from friends, neighbors, or fellow students, from parents or spouses, or from psychologists or other trained observers. In all these cases, the rater does not simply make a tally of behaviors, but instead interprets all the information he or she has about the person being rated—the target—as the basis for ascribing traits. Although friends and neighbors may be biased by favorable (or occasionally unfavorable) opinions of the target, they are less likely than the target to deceive themselves about the target's personality. Even more importantly, researchers who use peer ratings need not rely on any single view but can ask several raters for their judgments, examine the agreement among these, and combine them into an average rating.

Observer rating studies have played an important role in personality theory. The FFM itself was first identified in studies of ratings (Norman, 1963; Tupes & Christal, 1961/1992), and, as we will see, some important studies of age and personality have used peer, spouse, or expert ratings. In sheer quantity, however, self-report studies have far outnumbered observer rating studies, and perhaps the most important contribution of rating studies has been what they have been able to show about the validity of self-reports.

Rate yourself on the six Extraversion items listed earlier on a 1-to-5 scale, and then ask a friend who knows you well to rate you. You may be surprised at the differences of opinion. We would be guilty of failing to learn from decades of criticism if we did not acknowledge that there is a substantial element of illusion in the views we hold of others (Fiske, 1978) and in our idea of what their view of us is like (Swann, 1983). We easily and rapidly form opinions about others' personalities, often with little basis in fact. This is particularly true of those with whom we have only a passing acquaintance.

But these limitations should not cause us to despair. When knowledgeable raters are asked to assess a target's personality using reliable and valid trait measures, the results are generally encouraging (McCrae, 1982b). We asked subjects in the Baltimore Longitudinal Study of Aging (BLSA) to complete the self-report form of the NEO Personality Inventory; with their permission, we also obtained ratings from spouses

and from one to four peers (friends and neighbors), using an observer rating form of the NEO-PI (McCrae & Costa, 1989a). The first column of Table 3 gives intraclass correlations between pairs of peers who rated the same target. All the correlations are statistically significant, showing that ratings are not random impressions but instead are reflections of characteristics on which independent observers can agree. The extent of agreement is substantial for all five factors: The median value of .45 means that almost half of the variance in rating scores is attributable to the trait itself (Ozer, 1985), the other half being error—that is, mistakes and biases in perception, careless responding, legitimate differences of opinion, different frames of reference, and so on (McCrae, Stone, Fagan, & Costa, 1998). The second column of Table 3 shows that peers agree with spouses about as well as they agree among themselves. By averaging together two or more ratings we obtain even more trustworthy information about the target, as the errors of one observer are canceled out by opposing errors in another.

What about agreement between self-reports and ratings? Can we “see ourselves as others see us,” as Robert Burns wished? The last two columns of Table 3 suggest that we can, with substantial correlations of self-reports with averaged peer ratings and with spouse ratings. The highest correlations are seen when self-reports are compared with spouse ratings. Perhaps husbands and wives know the targets more intimately than do friends or neighbors, or perhaps they show higher agreement because couples discuss their personalities with each other and share their perceptions of themselves. In any case, this is one more

TABLE 3. Correlations between Peer Ratings, Spouse Ratings, and Self-Reports

NEO-PI factor	Agreement between			
	Peer–peer	Peer–spouse	Peer–self	Spouse–self
Neuroticism	.36	.45	.37	.53
Extraversion	.41	.26	.44	.53
Openness	.46	.37	.63	.59
Agreeableness	.45	.49	.57	.60
Conscientiousness	.45	.41	.49	.57

Note. All correlations are significant at $p < .001$; N 's = 144 to 719. Adapted from McCrae and Costa (1989a).

TABLE 4. Some Studies on Agreement between Self-Reports and Peer Ratings

Study	N	Participants	Raters	Measure	Average <i>r</i>
Amelang & Borkenau (1982)	321	Young adults	3	Adjective factors	.58
Kammann, Smith, Martin, & McQueen (1984)	60	Students	2	Adjectives	.30
S. V. Paunonen (cited in Jackson, 1984)	90	Students	1	PRF scales	.52
Marsh, Barnes, & Hocevar (1985)	151	Students	1	Self-concept scales	.58
John (1988)	50	Individuals	4	Adjectives	.31
Funder & Colvin (1988)	157	Students	2	CQS ^a items	.27
Costa, McCrae, & Dembroski (1989)	100	College men	4	NEO-PI domains	.51
Costa & McCrae (1992a)	142	Adults	2	NEO-PI domains	.43
Funder, Kolar, & Blackman (1995)	136	Students	1	NEO-PI domains	.36
Riemann, Angleitner, & Strelau (1997)	964	German adults	2	NEO-FFI scales	.55
Zawadzki, Strelau, Szczepaniak, & Sliwinska (1997)	548	Polish adults	2	NEO-FFI scales	.51

^aCalifornia Q-Set (discussed in a later section of this chapter).

piece of evidence that both ratings and self-reports are useful ways to measure personality, with at least as much truth as error.

Similar results and conclusions have been reported by many other researchers in recent years, as Table 4 shows. This research, conducted on a variety of samples ranging from American college students to Polish twins, firmly establishes the utility of both self-reports and ratings as methods of assessing personality. Table 4 also illustrates some other points. First, studies using scales or factors show higher agreement than those using single adjectives or items, because scales are more reliable than the individual items that compose them. Second, studies that average multiple raters tend to show higher correlations than those that rely on a single rater: In assessing personality, two (or more) heads are generally better than one.

Validity Scales

The studies in Table 4 begin in the 1980s. Until that time there had been relatively few studies comparing self-reports and observer ratings. Ratings are cumbersome to obtain, because raters must be identified, contacted, and persuaded to participate, and because some individuals may be unable or unwilling to provide the names of suitable raters. Researchers and clinicians have preferred the convenience of self-reports, even if they did not trust them. Consequently, a major tradition in personality assessment has sought ways to determine the validity of self-reports using special validity scales designed to outwit defensive or dishonest test takers, as a lawyer tries to trap a lying witness in a contradiction. Despite extraordinary efforts and ingenuity, however, it is far from clear that any of these scales really works—probably because most self-reports are reasonably honest to begin with.

Perhaps the most common strategy for creating validity scales was used by Edwards (1957), who had judges rate the social desirability of items on the Minnesota Multiphasic Personality Inventory (MMPI), and selected a set of highly evaluative items. People who wished to make a good impression would answer *yes* to the desirable items and *no* to the undesirable ones. By summing the number of such responses, test takers could be assigned a social desirability score. High scores were considered by many psychologists as grounds for throwing out the test. Another MMPI scale, the *K* scale, was designed to assess defensiveness, and scores on the psychopathology scales were “corrected” by adding a fraction of the *K* scale score (McKinley, Hathaway, & Meehl, 1948). In theory, this addition compensated for the defensive denial of psychological problems.

In practice, however, these approaches are very problematic. Logically, a high score on a social desirability scale might result from an attempt, conscious or unconscious, to make a good impression; but it might also result from honest answers by truly well-adjusted, agreeable, and conscientious people. It is desirable to be honest, and liars might falsely claim to be honest if asked—but truly honest people would, too! A claim of honesty is inherently ambiguous.

Crowne and Marlowe (1964) believed they had a solution to this problem: They wrote items that would almost never be literally true, such as “Before voting, I thoroughly investigate the qualifications of all candidates.” They reasoned that someone who answered *yes* to many

such items would clearly be lying. But test takers are not literalists; they are likely to say to themselves, “Surely these psychologists didn’t mean to ask if I actually study the voting records of every single political candidate, from President to dogcatcher. No one does, so that would be a stupid question to ask. What they must have meant to ask was whether I am a concerned citizen who takes voting seriously. Since I am and I do, I guess I should answer yes.”

Rational analyses of what subjects might be thinking when they respond will not suffice here; the only way to find out if validity scales succeed in identifying and correcting self-report biases is by empirical test. We can compare corrected and uncorrected self-reports with an impartial, external standard—say, observer ratings—to see whether correcting improves accuracy of self-reports. Several studies (Dicken, 1963; McCrae & Costa, 1983c; McCrae, Costa, Dahlstrom, et al., 1989; Piedmont, McCrae, Riemann, & Angleitner, 2000) convincingly show that the use of validity scales does nothing to improve the accuracy of self-reports in volunteer samples—indeed, it often impairs them. Most people apparently answer honestly, without any significant attempt to make themselves look good. It is possible that there is more defensiveness or socially desirable responding under other conditions, for example, when applying for a job (but see Ones, Viswesvaran, & Reiss, 1996, for a different view). But that possibility need not concern us here. The studies of personality and age that we examine all use volunteer subjects with no motivation to fake their scores, and we are justified in putting considerable faith in these self-reports as one line of evidence.

A QUESTIONNAIRE MEASURE: THE NEO PERSONALITY INVENTORY

Any of the instruments in Table 4 might be suitable for the study of aging and personality. In fact, however, most of our research has used some version of the NEO Personality Inventory, or NEO-PI (Costa & McCrae, 1985, 1989a), and a more detailed description of that instrument may give a better sense of the process of measurement and the meaning of the major dimensions of personality.

The Revised NEO-PI, or NEO-PI-R (Costa & McCrae, 1992a), consists of 240 items, phrased either in the first person for self-reports (e.g., “I really like most people I meet”) or in the third person for ob-

server ratings (e.g., “She has a very active imagination”). Respondents read each item and decide on a 1-to-5 scale if they *strongly disagree*, *disagree*, don’t know or are *neutral*, *agree*, or *strongly agree* with the statement. The responses are then summed to yield five basic domain scores for Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. When compared with data from normative samples that show how most people respond, these scores allow us to characterize the respondent. For example, an individual who scores higher than three-quarters of the normative sample on Agreeableness might be described as *compassionate*, *good-natured*, and *eager to cooperate and avoid conflict*. One who scored in the lower quarter of the distribution could be described as *hardheaded*, *skeptical*, *proud*, and *competitive*.

Table 1, in Chapter 1, listed some adjectives that describe low and high scorers on the five domains. In general, Neuroticism represents the proneness of the individual to experience unpleasant and disturbing emotions and to have corresponding disturbances in thoughts and actions (Vestre, 1984). Psychiatric patients who were traditionally diagnosed as neurotics tend to score very high on this dimension, but many individuals score high without having any psychiatric disorder: Neuroticism is a dimension of personality on which people vary only in degree. Extraversion is a more familiar term; it concerns differences in preference for social interaction and lively activity. Openness in our model should not be confused with self-disclosure (being willing to talk about one’s inner feelings); instead, it refers to a receptiveness to new ideas, approaches, and experiences (McCrae & Costa, 1997a). People who are closed to experience are not necessarily defensive, nor are they necessarily narrow minded in the sense of being judgmental and intolerant. Instead, they are characterized by a preference for the familiar, practical, and concrete, and a lack of interest in experience for its own sake.

Agreeableness is seen in selfless concern for others and in trusting and generous sentiments. Low Agreeableness (or Antagonism) is tough minded and hardheaded. Although agreeable people are nicer than antagonistic people, antagonism, too, has its virtues. We want surgeons to be ruthless in cutting out disease and lawyers to be aggressive in our defense. Conscientiousness is a dimension of individual differences in organization and achievement. Highly conscientious people are dutiful and self-disciplined, but also ambitious and hardworking, sometimes to

the point of being “workaholics.” Men and women low in Conscientiousness are more lackadaisical and easygoing and less exacting with themselves or others.

The NEO-PI-R was developed over a 15-year period, but in the first years we were concerned with only three factors, or *domains*—Neuroticism, Extraversion, and Openness (whence the name “NEO”). For these three domains we also developed measures of more specific traits, which we have called *facets* because they reflect specific sides or aspects of the broader domains. Each of the domain scales is composed of six subscales that measure these specific facets.

There is an important difference between the domains and the facets. Researchers coming from many different traditions have agreed on the importance of the five factors because these broad dimensions recur in most personality measures in one form or another. The finer distinctions within domains, however, are more arbitrary. For example, some theorists would combine Warmth and Gregariousness into a Sociability facet or break Anxiety down into Tension and Apprehension facets. There is certainly no reason why each domain should have exactly six facets. So far, however, no one has come up with a compelling theoretical or empirical basis for identifying facets, although most agree that some kind of facet system is necessary to address the subtleties of personality traits (e.g., Paunonen & Ashton, 2001). Our solution was to review the psychological literature and choose traits that seemed to cover the most important ways in which people differed. Although more than six distinctions could certainly be drawn in each domain, 30 constructs seemed to be pushing the limit for most users to grasp. Our system is a compromise between comprehensiveness and comprehensibility. After many years of use, it seems to be a reasonable approach to the measurement of facets (Costa & McCrae, 1995).

Facets of N, E, and O

Anxiety and *Angry Hostility*, the first two facets of Neuroticism, are the dispositional forms of two fundamental emotional states: fear and anger. Everyone experiences these emotions from time to time, but the frequency and intensity with which they are felt varies from one person to another. Individuals high in the trait of anxiety are nervous, high strung, and tense. They are prone to worry; they dwell on what might go wrong. Hostile people show a corresponding proneness to experi-

ence anger. They tend to be irritable and ill tempered, and may prove hard to get along with.

Two different emotions, sorrow and shame, form the basis of the facets of *Depression* and *Self-Consciousness*. As a trait, Depression is the disposition to experience sadness, hopelessness, and loneliness; depressed people often have feelings of guilt and of diminished self-worth. Individuals high in Self-Consciousness are more prone to the emotion of shame or embarrassment. They are particularly sensitive to ridicule and teasing, because they often feel inferior to others. (Self-Consciousness is often called shyness, but it should be noted that people who are shy in this respect do not necessarily avoid people. Introverts often claim to be shy when in fact they simply do not care to interact with others; they claim shyness as an excuse.)

Two of the facets of Neuroticism are more often manifest in behaviors than in emotional states. *Impulsiveness* is the tendency to give in to temptations and to be overwhelmed by desires. Because they have so little control (or perhaps because they experience such strong urges), impulsive people tend to overeat and overspend, to drink and smoke, gamble, and perhaps use drugs (Brooner, Schmidt, & Herbst, 1994). *Vulnerability* designates an inability to deal adequately with stress. Vulnerable people tend to panic in emergencies, to break down, and to become dependent on others for help.

Some individuals will be anxious but not hostile, or self-conscious but not impulsive. But in general, those high in Neuroticism are likely to be high in *each* of these traits. They are prone to violent and negative emotions that interfere with their ability to deal with their problems and to get along with others. We can envision the kind of chain of events such people may experience: In social situations they are anxious and embarrassed, and their frustration in dealing with others may make them hostile, further complicating matters. In compensation, they turn to the use of alcohol or food, and the long-term results are likely to be depressing. Although the emotions and impulses which disturb them may not occur simultaneously, they succeed one another with distressing regularity.

There is sometimes a tendency to ridicule people high in Neuroticism (think of the Woody Allen caricature) or to become impatient with their constant complaining. But a high level of Neuroticism is a very real burden that ultimately deserves compassionate understanding. (For a sensitive and insightful portrayal, see Michael Cunningham's [1998] depiction of Laura Brown in *The Hours*.)

The facets of Extraversion can be subdivided into three interpersonal and three temperamental traits. *Warmth*, or attachment, refers to a friendly, cordial, intimately involved style of personal interaction; by contrast, cold individuals are more likely to be formal and impersonal, with weak attachments to most other people. Warmth and *Gregariousness* (or the desire to be with other people) together make up what is sometimes called sociability. Gregarious people like crowds; they seem to relish sheer quantity of social stimulation. *Assertiveness* is a third facet of Extraversion; assertive people are natural leaders, easily taking charge, making up their own minds, and readily expressing their feelings and desires.

The three facets of Extraversion we have called temperamental are *Activity*, *Excitement Seeking*, and *Positive Emotions*. Extraverts like to keep busy, acting vigorously and talking rapidly; they are energetic and forceful. They also prefer environments that stimulate them, often going in search of excitement. Fast cars, flashy clothes, risky undertakings hold an attraction for them. The active and exciting life of extraverts is reflected emotionally in the experience of positive emotions. Joy, delight, zest, and jocularly are part of the package of traits in the domain of Extraversion. Once again, all these dispositions are synergistic, working together to form a personality syndrome. Activity leads to excitement and excitement to happiness; the happy person finds others easier to get along with, and congeniality easily turns to leadership.

We measure Openness to Experience in six different areas. Openness in *Fantasy* refers to a vivid imagination and a tendency to develop elaborate daydreams; in *Aesthetics* it is seen in sensitivity to art and beauty. Aesthetic experience is perhaps the epitome of Openness, since it is pure experience for its own sake; we know from studies of occupational interests that a preference for artistic activities is especially characteristic of open people (Costa, McCrae, & Holland, 1984). As Carl Rogers might have predicted, open individuals experience their own *Feelings* strongly, and they value the experience, seeing it as a source of meaning in life.

Openness to *Actions* is the opposite of rigidity: Open people are willing to try a new dish or see a new movie or travel to a foreign country. Openness to *Ideas* and *Values* are also facets of the domain. Open people are curious and value knowledge for its own sake. Perhaps because they are willing to think of different possibilities and to empathize

with others in other circumstances, they tend to be liberal in values, admitting that what is right and wrong for one person may not be applicable in other circumstances.

Facets of A and C

The original NEO-PI did not measure specific facets of Agreeableness and Conscientiousness; only global domain scales were included. But the NEO-PI-R (Costa & McCrae, 1992a; Costa, McCrae, & Dye, 1991) was expanded to include six facets for A and C. By now there has been considerable research on these facets, including a 6-year longitudinal study (Costa, Herbst, McCrae, & Siegler, 2000) that we consider in a later chapter.

Agreeable people are trusting, believing the best of others and rarely suspecting hidden intents. Age changes in *Trust* would be of particular interest because Erikson (1950) regarded this as the earliest and most fundamental outcome of psychosocial development. According to his views, those who do not develop trust can never really advance toward industry, identity, and intimacy. Just as agreeable people trust others, so they are themselves trustworthy, marked by candor and *Straightforwardness*. The selflessness of agreeable people is seen in their considerateness and desire to help others—a facet we call *Altruism*. Agreeable individuals are meek, deferring to others rather than aggressively pushing for their own ends. This facet is called *Compliance*, a nod toward *Friendly Compliance*, an old label for the Agreeableness domain (Digman & Takemoto-Chock, 1981). Agreeable people are humble, showing *Modesty* in their assessment of their own abilities and importance. Low scorers on this facet might be considered narcissistic. Attitudinally, agreeable people exhibit *Tender-mindedness* and sentimentality, and may be an easy touch for charities and good causes.

Conscientious people are rational, informed, and generally think of themselves as being high in *Competence*. Part of their success results from their organization and *Order*, which makes them efficient in work. In some respects, conscientious people are inhibited, adhering scrupulously to their moral precepts; like Admiral Horatio Nelson, they have a strong sense of *Dutifulness* (Costa & McCrae, 1998a). They are high in *Achievement Striving*, pursuing excellence in everything they do; and they are necessarily high in *Self-Discipline* to be able to accomplish their goals. Finally, they are characterized by *Deliberation*, making plans in

advance and thinking carefully before acting. Theirs is a life clearly directed along the paths they choose to pursue.

Making Distinctions

It is not difficult to see the relationships among traits within a single domain. But it can be difficult to make the distinctions between and among domains that the facts demand. Human beings have a tendency to evaluate ideas and organize them in terms of good and bad. It is easy to see Neuroticism as bad and Extraversion, Openness, Agreeableness, and Conscientiousness as good, and then to assume that individuals high in Neuroticism are likely to be introverted, closed, antagonistic, and unreliable. It doesn't work that way. The five factors are independent; consequently, an individual high in Neuroticism is as likely to be introverted as extraverted, closed as open. Five dimensions cannot be collapsed into one without serious distortions.

One way to make the necessary conceptual distinctions is to concentrate on the opposite poles, the low ends, of the traits we are discussing. Introverts are cold but not hostile; loners but not self-conscious; not particularly cheerful, but not necessarily depressed either. Individuals who are closed to experience are not necessarily maladjusted; in fact, they are somewhat less likely to experience violent emotions by virtue of their lesser sensitivity to emotions of all kinds. They may not care to travel to strange places and meet unusual people, but they are just as likely as open people to seek excitement and crowds of people—as long as both are of the familiar variety.

There is another mistake in contrasting “good” and “bad” dimensions of personality: We may be mistaken about which is which. From the individual's point of view, there are distinct advantages to being well-adjusted and extraverted. Such people usually report themselves to be the happiest, most satisfied with life (Costa & McCrae, 1980a). But all traits have passed the evolutionary test of survival, and from society's point of view all kinds of people are necessary: those who work well with others and those who can finish a task on their own; those who come up with creative new ways of doing things and those who maintain the best solutions of the past. There are probably even advantages to found in Neuroticism, since a society of extremely easygoing individuals might not compete well with other societies of suspicious and hostile individuals. Cultures need members fit for war as well as peace,

work as well as play, and some theorists (Hogan, 1979) have even suggested that such societal advantages may account for individual differences in an evolutionary sense: Human beings have evolved consistent differences because this range of characteristics is useful in the survival of the social group.

THE COMPREHENSIVENESS OF THE FIVE-FACTOR MODEL

We know that the FFM encompasses the lay conception of personality as embodied in the trait names of the English language and that the NEO-PI-R provides both self-report and observer rating measures of the factors (McCrae & Costa, 1985b, 1987). We can now turn to the question of whether the FFM is truly comprehensive: Can it account for the traits identified by other psychological systems? A series of studies using the NEO-PI along with other instruments points to an affirmative answer.

Perhaps the most impressive evidence of comprehensiveness comes from analyses of the California Q-Set, or CQS, developed by several panels of psychologists and psychiatrists who explicitly tried to include every important aspect of personality suggested by their psychodynamic orientation (J. Block, 1961). We know that their selection was not biased by the FFM, because they had completed their work before it was first reported (Tupes & Christal, 1961/1992).

In our study, we had 403 men and women from the BLSA describe themselves using the CQS; when we analyzed the 100 items, we found five factors whose defining items are given in Table 5. It was clear to us that these were essentially the same factors that we measured in the NEO-PI, and correlations between the CQS factors and NEO-PI scales confirmed this hypothesis. The FFM organizes CQS items as well as it organizes trait names.

One particularly elegant way to demonstrate convergences between two instruments is by factoring the scales from both instruments together in a joint analysis. When Wiggins's revised Interpersonal Adjective Scales (IAS-R; Wiggins, Trapnell, & Phillips, 1988) were factored along with NEO-PI factors, all the IAS-R scales showed their highest loading on the Extraversion and Agreeableness factors (McCrae & Costa, 1989d). As Figure 1 in Chapter 2 shows, the traits of the interpersonal circle can be arranged around these two dimensions. From this

TABLE 5. California Q-Set Items Defining the Five Factors

Factor	Low scorer	High scorer
Neuroticism	Calm, relaxed Satisfied with self Clear-cut personality Prides self on objectivity	Thin-skinned Basically anxious Irritable Guilt-prone
Extraversion	Emotionally bland Avoids close relationships Overcontrol of impulses Submissive	Talkative Gregarious Socially poised Behaves assertively
Openness	Favors conservative values Judges in conventional terms Uncomfortable with complexities Moralistic	Values intellectual matters Rebellious, nonconforming Unusual thought processes Introspective
Agreeableness	Critical, skeptical Shows condescending behavior Tries to push limits Expresses hostility directly	Sympathetic, considerate Warm, compassionate Arouses liking Behaves in a giving way
Conscientiousness	Eroticizes situations Unable to delay gratification Self-indulgent Engages in fantasy, daydreams	Behaves ethically Dependable, responsible Productive Has high aspiration level

Note. Adapted from McCrae, Costa, and Busch (1986).

diagram it appears that the traditional axes labeled “Love” and “Status” are intermediate between E and A. In other words, both systems describe the same circle of traits, but they do so from somewhat different starting points.

Joint factor analysis (Costa & McCrae, 1988a) was also useful in understanding Murray’s (1938) needs as measured by Jackson’s (1984) Personality Research Form (PRF). Table 6 gives factor loadings; by concentrating on the larger loadings (in boldface), the relationship between needs and personality factors can be seen. These data (from a group of people originally recruited to provide ratings on our BLSA subjects) suggest that individuals high in Neuroticism worry about others’ opinions of them (need for Social Recognition), are defensive and guarded (need for Defence), and want care and sympathy (need for Succorance). Extraverts need social contact (Affiliation), attention (Exhibi-

tion), and fun (Play). Open individuals appreciate variety (Change), intellectual stimulation (Understanding), and aesthetic experience (Sentience), and are adventurous (low Harmavoidance) and unconventional (Autonomy). Individuals high in Agreeableness enjoy helping others (Nurturance) and tend to be self-effacing and modest (Abasement); by contrast, antagonistic people are domineering (Dominance) and quarrelsome (Aggression). Those who score high in Conscientiousness value organization (Order) and accomplishment (Achievement) and are

TABLE 6. Joint Factor Analysis of NEO-PI Factors and PRF Scales

Variable	Varimax-rotated principal component				
	N	E	O	A	C
NEO-PI factor					
Neuroticism	81	-17	02	09	-14
Extraversion	13	83	13	-03	13
Openness	-09	-07	78	04	-11
Agreeableness	-01	15	-18	72	-23
Conscientiousness	08	-15	00	23	77
PRF need					
Abasement	06	-14	12	58	08
Achievement	05	03	46	02	64
Affiliation	04	83	-13	19	11
Aggression	43	07	14	- 68	-21
Autonomy	- 42	-33	47	-26	-10
Change	-06	21	60	-12	-11
Cognitive Structure	19	07	-23	-30	53
Defendence	53	-07	-13	- 48	-05
Dominance	00	38	45	-46	32
Endurance	-16	07	33	15	52
Exhibition	05	65	23	-31	-03
Harm avoidance	21	05	- 52	32	09
Impulsivity	29	11	24	03	- 61
Nurturance	25	49	10	55	06
Order	-05	12	-25	-17	64
Play	-13	65	07	-06	-37
Sentience	11	29	53	13	-09
Social Recognition	60	34	-10	-19	10
Succorance	53	40	-34	18	-14
Understanding	-02	00	64	10	16

Note. Decimal points omitted. Loadings greater than .40 in absolute magnitude are given in **boldface**. Adapted from Costa and McCrae (1988a).

persistent (Endurance), careful (low Impulsiveness), and deliberate (need for Cognitive Structure).

Table 6 also shows something else: Many of the needs have high loadings on two or more factors. Nurturance, for example, is related to both Extraversion and Agreeableness; Aggression is related to both Neuroticism and (low) Agreeableness. These findings are completely reasonable (McCrae & Costa, 1989d). We expect nurturant people to be both sociable and selfless and aggressive people to be both angry and selfish. But joint loadings like this make the correspondence between needs and the five factors less obvious. Researchers who begin by thinking in terms of nurturance and aggression are unlikely to come up with the dimensions of Extraversion, Agreeableness, and Neuroticism, even though these dimensions make perfect sense in hindsight. Perhaps this is one of the reasons why the FFM was not discovered sooner.

Similar analyses have shown that the scales in the Guilford–Zimmerman Temperament Survey (GZTS), the Eysenck Personality Questionnaire (EPQ), the Myers–Briggs Type Indicator (MBTI), and the California Psychological Inventory (CPI) can also be related to the basic five factors (McCrae, 1989; McCrae, Costa, & Piedmont, 1993). Table 7 shows how the scales can be classified. Note that the EPQ contains no measure of Openness, the MBTI no measure of Neuroticism, and the GZTS a disproportionate number of scales measuring Neuroticism and Extraversion. Only the broad summary measures from the CPI are given, but some of the individual scales do measure Openness (e.g., Achievement via Independence). Curiously, Agreeableness is largely absent from CPI content. The five factors appear to be both necessary and sufficient for describing the basic dimensions of personality; no other system is as complete and yet as parsimonious.

Researchers have related almost every major personality inventory to the NEO-PI-R, including the Adjective Check List (Piedmont, McCrae, & Costa, 1991), the Edwards Personal Preference Schedule (Piedmont, McCrae, & Costa, 1992), the Comrey Personality Scales (Hahn & Comrey, 1994), Lorr's Interpersonal Style Inventory (McCrae & Costa, 1994), the MMPI (Costa, Busch, Zonderman, & McCrae, 1986), and the Sixteen Personality Factor Questionnaire (16 PF; Conn & Rieke, 1994). Indeed, in an *Annual Review* article on personality structure, Ozer and Riese concluded: "Just as latitude and longitude permit the precise description of any location on earth, the FFM prom-

TABLE 7. Classification of Traits from Some Alternative Systems

System	N	E	O	A	C
GZTS		General Activity –Restraint Ascendancy Sociability			Restraint
	–Emotional Stability –Objectivity		Thoughtfulness		
	–Friendliness –Personal Relations –Masculinity			Friendliness	
EPQ	Neuroticism	Extraversion		–Psychoticism	–Psychoticism
MBTI		Extraversion	Intuition	Feeling	Judgment
CPI		–Internality			Norm-favoring
	–Realization				

Note. Minus signs indicate that the scale measures the opposite of the factor.

ises the hope of similarly locating personality dispositions” (1994, p. 361).

Finally, considerable research indicates that the Axis II diagnostic categories for personality disorders may also be classifiable in terms of the five factors (Costa & Widiger, 2002). In one of the first of these studies, Wiggins and Pincus (1989) showed that scales which measure schizoid disorders are closely related to Introversion whereas those measuring histrionic disorders are related to Extraversion. Borderline disorders are characterized by high Neuroticism; dependent disorders, by high Agreeableness. Individuals high in Conscientiousness appear to be prone to compulsive disorders. Only Openness is not clearly implicated in the current list of personality disorders, and perhaps future research will suggest that there are also pathological forms of Openness and Closedness. The FFM appears to encompass pathological as well as normal personality traits.

* * *

The meaningfulness of trait psychology and the validity of trait measures are crucial to all the other arguments about aging and personality we offer in later chapters, because the great majority of empirical studies of aging and personality employ these kinds of measures. In Chapter 8 we return to the theories of personality to consider alternative conceptions of personality and its assessment. In the meantime we will assume that we have provided adequate justification for looking seriously at what happens to traits over the adult life course. We have argued that traits are indispensable for laypeople and for psychological theorists; that there is growing agreement about how many important groups of traits there are; and that there is encouraging evidence that several ways of measuring personality all lead to similar descriptions of people and thus are dependable bases for inferences about aging and personality. We can now turn to the data on aging. What effects do aging and the accumulation of experience have on personality traits?

The Search for Growth or Decline in Personality

If we want to know whether personality changes across the adult years, the most straightforward way of proceeding would seem to be to measure personality in young and old people and see if there are any noticeable differences. This simple prescription turns out to be more complex than one might at first imagine. There are several excellent discussions of measuring change and disentangling aging, birth cohort, and time-of-measurement effects (e.g., Kausler, 1982; Schaie, 1977), complete with diagrams and statistical formulas. Our purpose here is not to teach methodology but to show why a question as apparently straightforward as “Does personality change with age?” is not straightforward at all and to review evidence on whether personality grows or declines with age.

Growth and decline are the most obvious and fundamental of developmental processes. Children grow taller each year until they reach adult stature; thereafter they remain at about the same height for many years, until in old age the settling of bones in the vertebral column makes them shrink somewhat. Many biological functions, notably sexuality, show similar curves, as do some, though by no means all, of the intellectual functions, such as short-term memory and spatial visualization abilities.

Many physical, biological, and social functions, however, do not exhibit the pattern of growth and decline. Some characteristics, like gender or eye color, are established by birth and show no significant change across the lifespan. Variables of another class go through a pe-

riod of growth but show little or no decline. Vocabulary, for example, tends to grow dramatically in childhood and then to level out after the end of formal education, with little change thereafter.

These trends do not exhaust the possibilities. Perhaps growth in personality is unlimited, like the physical growth of some fish. Perhaps the pattern of change is not a simple up or down but a complicated pattern of peaks and troughs—a possibility suggested by some of the stage theories we have mentioned. Rather than belabor the possible, let us turn to the facts and see if we can simply *describe* the age trends in personality dimensions. Later we can worry about what they mean or where they come from.

CROSS-SECTIONAL STUDIES OF PERSONALITY DIFFERENCES

There are some fairly obvious ways of going about looking for this kind of age trend. First, we can measure an individual's personality now, and then return in 10, or 20, or 40 years to measure it again. We could then tell what net change, if any, had occurred in the meantime for each individual. This is a longitudinal design, in many respects the best way to study aging. The major problem with longitudinal studies is quite simple: They take too long. Few scientists are willing to wait 10 years for results, let alone 40. Gardeners may be willing to plant trees now that will not bear fruit for 10 years, but few researchers see themselves as gardeners. In fairness, there are some differences. Gardeners know that if they wait patiently they will be rewarded with apples or cherries or plums. Researchers never know what they will find out, or if the results will turn out to be interesting or informative. If they knew the answers, they would not need to ask the questions.

What one really wants is some shortcut method of answering the question of age changes, and the cross-sectional method is invaluable as that kind of tool. In cross-sectional designs, one takes a group of younger people and a group of older people and compares them on the traits of interest. We might observe that 40% of men over 60 were bald whereas only 2% of men under 30 were. Hence, it would seem, baldness is age related. This is an example of the kind of cross-sectional observation we all make, and it is almost certainly a correct conclusion.

Studies that compare groups of different ages are called *cross-*

sectional because they take a slice at one time of individuals of various ages. Note that this strategy has a number of advantages. For example, if we measure enough people, we can look not only at young versus old but at much finer groupings: We can contrast people 38–43 years old with people 44–49 years old to see if the former are showing signs of a midlife crisis. We can search for a curve of growth and decline by decade or year to say at exactly what point there is a peak in creativity or productivity or depression. In a matter of weeks or months a researcher can estimate effects of age on whatever variable he or she is interested in studying.

By this method, thousands of studies on age and personality have been conducted and hundreds published in the scientific literature. In 1977, Neugarten reviewed the findings that had been reported: egocentrism, dependency, introversion, dogmatism, rigidity, cautiousness, conformity, ego strength, risk taking, need for achievement, locus of control, creativity, hope, the self-concept, social responsibility, morale, dreaming, and attitudes toward aging were all found to be different in different groups. However, the studies rarely agreed with each other: Some researchers found that older people are more dogmatic, some that they are less. (Doubtless other researchers have found no age differences at all, but it is sometimes difficult to find a journal that will publish such “uninteresting” results.) Only introversion, according to Neugarten, seems to show a consistent pattern of increase in the latter half of life.

In updating Neugarten’s review we can take advantage of the Five-Factor Model (FFM) to organize results, and we can test the replicability of findings by comparing age effects for different scales that measure each of the five domains. We will begin with three large-scale studies that use instruments that cover all or most of the five domains. Siegler, George, and Okun (1979) studied a sample of 331 men and women initially aged 54–70 who completed the 16PF. Douglas and Arenberg (1978) examined age differences on the GZTS among two groups of men in BLSA: 605 men tested between 1958 and 1968, and 310 men tested between 1968 and 1974. Subjects ranged in age from 17 to 98, so age differences anywhere in adulthood should be apparent. In our study of the PRF (Costa & McCrae, 1988a) we examined age correlates in a sample of 296 men and women aged 22–90 who were friends and neighbors of BLSA participants. (A significant positive correlation

of scale with age would suggest that the trait increases with age; a negative correlation suggests that the trait declines.)

Four scales in the 16PF—low Emotional Stability, Suspiciousness, Tension, and Guilt-proneness—measure aspects of Neuroticism; none of them showed cross-sectional age differences. Four GZTS scales—Emotional Stability, Objectivity, Personal Relations, and Masculinity—tap emotional adjustment, the polar opposite of Neuroticism. Masculinity was slightly lower for older subjects in both samples studied by Douglas and Arenberg (1978), whereas Personal Relations was slightly higher; there were no differences in Emotional Stability or Objectivity in either sample. Age was also uncorrelated with PRF scales related to Neuroticism: Social Recognition, Defence, and Succorance. Apparently there are few age differences in Neuroticism.

There is more evidence of age differences in Extraversion. GZTS General Activity and Ascendancy were lower in older BLSA men in both samples, as was Sociability in the first BLSA sample. Age was also negatively related to PRF needs for Exhibition and Play. However, there were no age differences in the need for Affiliation, nor in any of the 16PF Extraversion scales: Warmth, Assertiveness, Surgency, Adventurousness, and Group Dependence.

Age trends in Openness are mixed. Needs for Change and Sentience appeared to decline with age, whereas the need for Understanding increased. GZTS Thoughtfulness also increased, but there were no age differences in needs for Harmavoidance or Autonomy or in 16PF scales measuring aesthetic sensitivity, imaginativeness, and liberal thinking.

Agreeableness is not well measured by the 16PF scales, but GZTS Friendliness was a bit higher among older subjects, as was the need for Abasement. Needs for Dominance and Aggression appeared to decline with age. All these data point to a cross-sectional increase in Agreeableness.

Although GZTS Restraint was higher for older men in both of Douglas and Arenberg's samples, other measures of Conscientiousness—16PF Superego Strength and Self-Control and PRF needs for Achievement, Order, Endurance, Cognitive Structure, and low Impulsiveness—showed no relationship to age.

There are several different ways to interpret these studies, and most of the interpretations would be championed by one writer or another.

One reasonable reading of this literature would be that there are age-related declines in some aspects of Extraversion and increases in Agreeableness but that most other aspects of personality remain constant across the lifespan. But many researchers in the field would be skeptical about this conclusion. They would point out that cross-sectional findings may result from many causes besides maturation. One alternative interpretation is provided by sampling bias.

Sampling Bias

Virtually all psychological research makes use of sampling and statistical inference in reaching its conclusions. We do not give personality tests to all 80-year-olds and all 30-year-olds; we give them to some few whom we can persuade to take them. Yet the conclusions we are interested in are not limited to the group we happen to have measured but concern aging in general or 80-year-olds in general. A great deal of acute mathematical thinking has gone into the development of statistical tests that allow us to determine whether the differences we see between two random samples are real or simply a chance outcome of the particular group of people we happened to examine. In general, the larger the sample, the better the estimate of the true score for the population, the whole group about which we want to make generalizations. Samples of about 2,000 are routinely used to estimate the feelings of 250 million Americans on such issues as the energy crisis or the likelihood of an economic depression.

Sampling itself, then, is not a problem for science. The trouble comes in the requirement that the sample be *random*. If we could go through the records of the U.S. Census Bureau, selecting every 100,000th citizen, we would be in a good position to make the best of a cross-sectional study. But samples tend to be quite different in real life. In practice, we find a sample by asking the cooperation of a senior citizens' group, or advertising in the paper, or asking our friends if they or their parents would participate in our study. This is sometimes called *haphazard sampling*, since it is neither really random nor exactly systematic.

In the early days of gerontological research, some dreadfully flawed conclusions were reached because of this problem. Old people were selected from nursing homes, since there were few senior citizens' clubs in those days; young people, of course, came from the colleges at which the researchers worked. The so-called age differences that emerged from

these studies painted a dismal picture of old age: The older people were rigid, depressed, less intelligent, less mentally healthy, and so on. The fact was somehow forgotten that college students represented the elite of their generation whereas nursing home residents were by and large the worst off of theirs. If we were to extrapolate from these cross-sectional findings, we would discover, for example, that by the time we are in our 80s we will only have had an 8th-grade education! Errors this obvious are rarely made today. But sampling still presents problems at whose magnitude we can often only guess.

Part of the problem stems from the fact that researchers tend to be specialists. Personality psychologists are expert at constructing scales but tend to be less concerned with sampling. Sociologists and epidemiologists are meticulous about their sampling strategies but may have limited knowledge about psychometrics and are often constrained by the time limitations imposed by the interview format they favor. Rarely does a gerontologist combine both sets of skills in looking for age differences, so real progress often depends on collaborative work among many scientists from different disciplines.

A good example is the National Health and Nutrition Examination Survey I Follow-up Study, or NHANES (Cornoni-Huntley et al., 1983). Originally designed to provide basic information on Americans' health, this survey was made the basis for a nationally representative study of aging. More than 10,000 adults ranging in age from 35 to 85 were surveyed between 1981 and 1984; included in the interviews were short scales to measure Neuroticism, Extraversion, and Openness (Costa & McCrae, 1986b).

Figure 2 summarizes the results of cross-sectional analyses for Black and White men and women. It is obvious that the lines are essentially flat, showing only tiny declines across five decades. (Because of the huge sample size, even these slight effects are statistically significant.) Data from this representative sample replicate other cross-sectional findings of lower Extraversion among older subjects, but they also underscore the limited magnitude of the age differences. It would seem a little foolish to concoct a theory of age and personality around the effects seen in Figure 2. But we have not yet heard the last of them.

Agreeableness and Conscientiousness were not included in the NHANES study, so we do not know what pattern they would have shown in a representative sample. Even if the cross-sectional increases in Agreeableness we reviewed earlier had been replicated, however, it

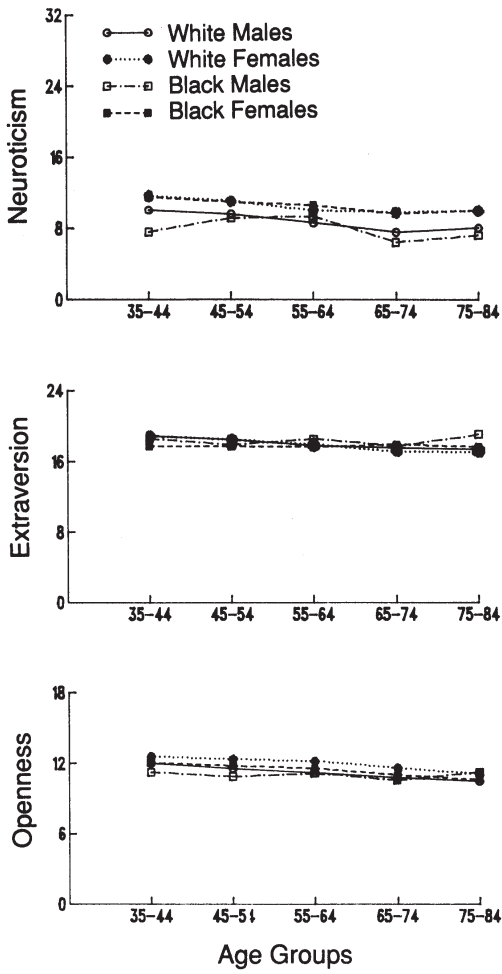


FIGURE 2. Mean levels of Neuroticism, Extraversion, and Openness to Experience for 10-year age groups of White men, Black men, White women, and Black women, aged from 35 to 84 years. Adapted from Costa, McCrae, et al. (1986).

could be argued that the effect was the result of another artifact, selective mortality. There are preliminary but intriguing data suggesting that antagonistic hostility is associated with heart disease (Costa et al., 1989): Young and middle-aged men who are very low in Agreeableness appear to be at higher risk for experiencing a heart attack. It is possible, therefore, that Agreeableness is higher among older groups only be-

cause the most disagreeable members of their generation have already died. Agreeableness itself need not increase with age in order to find a cross-sectional age difference.

Cohort Effects

Spurious age differences in cross-sectional studies can also result from generational or birth-cohort effects. When we make cross-sectional comparisons, we are in effect assuming that the young people of today will eventually come to resemble the current generation of older people. In some respects, such as the increased incidence of baldness, they almost certainly will. But we do not know whether they will also have the same attitudes, same habits, or the same personality as today's elderly. In short, we do not know whether personality is shaped more by aging or by cohort factors, by intrinsic maturation or by the social history of specific generations of people.

In the early days of personality theory, under the influence of Freud's models of development, it was assumed that personality was formed early in life as the result of a complex series of interactions between child and parents. Overly rigorous forms of toilet training and general discipline were held to be the source of fixations, inhibitions, and neuroses in later life. Whereas Freud himself never advocated such a simple relationship, many psychologists, pediatricians, and educators came to the conclusion that we would all be more emotionally healthy if we were given more freedom as children. Dr. Benjamin Spock's (1946) immensely influential writings on child care took a step in this direction, and books such as Neill's (1977) *Summerhill* showed what childhood could be like in an atmosphere of freedom. Over the past half-century, corporal punishment has almost disappeared from schools and has certainly declined in most households. Indeed, many psychologists and educators today are calling for a return to firmer discipline, just as Neill tried to distinguish between "freedom" and "license."

The point is not that today's form of discipline is better or worse than that enjoyed by children 50 years ago—the point is that it is demonstrably different. According to the basic tenets of some forms of personality theory, this difference in the treatment of children ought to have the most dramatic consequences for their adult personality. If child rearing changes across generations and if adult personality depends on child rearing, different generations should show major differences in

personality. These differences, however, would reflect not the action of aging but the accident of birth into a particular generation.

This confusion of generational or birth-cohort differences with age changes is not merely hypothetical. There are a number of documented cases in which cohort differences have been shown to be responsible for cross-sectional findings. Bones in the leg, for example, are typically a bit shorter in older men than in younger men. This is not a maturational effect, however, but a cohort effect (Friedlander et al., 1977): More recent generations of children have had better diets and have as a result grown taller in childhood and adolescence. Somewhat closer to the topic at hand, generational differences have also been shown to account for most of the age differences in vocabulary that have been reported (Schaie & Labouvie-Vief, 1974). Most vocabulary is learned in formal schooling, and today's adults had much more of that than their grandparents did.

If we return to the cross-sectional studies we have reviewed, we may now want to revise our conclusions. Perhaps the consistent decline in Extraversion is the result of changing styles of child rearing. When children were taught to speak only when spoken to, they may have developed lifelong habits of reticence. The formal interactions of children with their parents may have put a permanent damper on interpersonal warmth and spontaneity. Thus, the older generation we see now may be more introverted not because of age but because of what passed for good breeding a few decades ago. Today's outspoken, uninhibited children may turn out to be outspoken, uninhibited oldsters half a century from now.

On the other hand, it is also possible that very different processes are going on under the camouflage of age differences. If there are such significant changes in child rearing, shouldn't we be seeing many and massive age differences when we compare distant generations? Or is it possible that the absence of conspicuous age differences is the result not of stability but of equal and opposite effects of aging and cohort? This perplexing idea, which plays a central role in many arguments about aging research, deserves a bit more discussion.

Suppose that, as progressives hope, the change in child-rearing practices over the past 50 years has really had the beneficial effects that it was supposed to: Suppose that today's young adults are mentally and emotionally healthier than the young adults of 1930. Suppose further that, in the process of living and learning how to adapt, all people tend

to grow in mental health with the years and to overcome the handicaps of a less-than-optimal upbringing. Then we would say that society as a whole is getting healthier and healthier as each new generation starts off better and as each individual, every day in every way, gets better and better. Note that under this optimistic set of assumptions, there would be little evidence of age differences when cross-sectional comparisons were made: The older individuals would have risen from a level of poor mental health to good mental health, but they would not now differ from young adults whose good mental health is the result of enlightened child-rearing practices. Thus two major effects on well-being—one linked to socialization, one to aging—could jointly give the appearance of “no difference.”

Lest the reader be unduly encouraged by this hypothesis, note that the opposite set of premises could yield the same conclusion. Perhaps what we have seen in the past decades is the decline of civilization, a breakdown of discipline that leaves each successive generation with fewer and weaker inner resources. If at the same time aging has the same deleterious effects on mental health as on physical health, then all of us are getting worse with each passing day. And yet, again, a comparison of young with old people would show equal levels of mental health—or, in this case, mental illness.

Alternative interpretations like these are the despair of social scientists and have prompted many to abandon altogether the cross-sectional study. These researchers prefer to study aging as an ongoing process in individuals. They conduct longitudinal research.

LONGITUDINAL DESIGNS: TRACKING CHANGES OVER TIME

Most of the problems associated with cross-sectional studies can be solved by using the other major method of aging research, longitudinal designs. In these, a group of individuals is selected and measured at one time, then followed and measured repeatedly over a period of years, perhaps for decades. In contrast to the age differences estimated by cross-sectional studies, longitudinal studies estimate age changes. That is, they can see directly how the individuals under study have changed and perhaps at what point the changes occur. (As we will see, this does not necessarily mean that the changes are *caused* by aging.)

The major appeal of the longitudinal design is its ability to separate changes from generational effects. People only live once. In their youth they may absorb the characteristic tone of their generation, but having done so, they retain it for life. Longitudinal changes cannot be the direct result of generational differences, nor is there much evidence to suggest that individuals of different generations change in different ways or at different rates, at least with regard to personality.

In contrast to the hundreds of cross-sectional studies, there are only a relative handful of longitudinal studies of adult personality. The reasons should be clear: It is extremely difficult to keep track of people over a period of many years, and the rewards of longitudinal research are remote. In fact, virtually all the studies to be reported demonstrate the altruism of a set of foresighted scientists who began the measurement of personality in studies others would eventually report.

These studies vary considerably in the kinds of individuals studied, the methods and instruments used to measure personality, and the intervals at which measurements were made. They are quite consistent, however, in their conclusions: There is little evidence of substantial longitudinal change in personality characteristics in the period of adulthood, from 30 to 80 years of age. Let us turn to some of the specifics. We will review longitudinal analyses of the I6PF because the results are typical and because the existence of two independent studies makes informative comparisons possible. Later in this chapter we return to more recent studies.

Stability in the I6PF

Our early research on aging and personality was conducted in conjunction with a longitudinal program sponsored by the Veterans Administration (VA) in Boston, the Normative Aging Study (Costa & McCrae, 1978). The participants in this project were 2,000 male veterans, ranging in age from 25 to 90 years. Most were White and long-term residents of the Boston area. Although there were relatively few individuals from the lowest socioeconomic levels, there was a reasonable representation of subjects with a high school education and beyond. Of course, as in all such studies, the participants were volunteers; in this case they had volunteered to return to the VA Outpatient Clinic every 5 years for tests. They also completed questionnaires mailed to them at

home as the need arose. In addition to medical examinations, the men were given occasional assessments of personality and cognitive abilities. Between 1965 and 1967 about half the participants were given Cattell's 16PF.

In 1975, we readministered a form of the 16PF to a group of 139 men, originally aged 25–82. Among other things we were interested in seeing if there were systematic declines or increases in the traits measured by the 16PF. Of the 16 scales, two, intelligence and social independence, were significantly higher at the second administration; 14 showed no evidence of change over the 10-year interval studied. Tension, adventurousness, liberal thinking, tender-mindedness, and super-ego strength—among other traits—neither increased nor decreased for the average man in this sample. In retrospect, we view this as the major finding of the study. At the time, however, we were as curious as everyone else about the changes. The scales that absorbed our attention were the two that changed. But as we examined our findings in detail, it became clear that they were probably *not* the elusive age changes we had been seeking.

In 1975 the participants scored higher on measures of intellectual brightness and social independence. We suspect the increase in intelligence scores was probably a matter of testing. When people are given the same test twice, they tend to do better the second time, even though they are not actually any more intelligent—this is known as a *practice effect*. In addition, the test was given at the laboratory on the first occasion, under timed conditions; in 1975, the men completed the test at home, at their own pace. Allowing more time may have improved scores. We concluded that it was unlikely that the men had actually become more intelligent with age. (In any case, we do not regard intelligence as a dimension of personality, and we leave it to our colleagues in cognitive aging to chart the course of mental abilities in adulthood.)

The longitudinal change in the social independence scale was more puzzling. For one thing, there were no cross-sectional differences in this trait. That probably meant that, if real longitudinal changes were taking place, they were not showing up in cross-sectional comparisons because they were somehow obscured by generational differences. It is precisely because of the possibility of such confusions that researchers turn to longitudinal designs; when they find evidence of them, however, it is usually an unpleasant surprise. We were faced with the prospect of explaining not only why individuals became less dependent on the ap-

proval of social groups as they aged but also why different generations (or, as we will see, times of measurement) also influenced this variable.

Before exerting our creativity in looking for an explanation of this curious finding, we should be sure there is really something here to explain. The fundamental principle of science is reproducibility: A phenomenon must be dependable, regardless of how or by whom it is observed. Few phenomena in the psychological sciences show the invariance of the laws of physics, because our constructs are more abstract and less easily and accurately measured. But we have a right to demand that results be generally replicable—that most investigators will report the same general findings and that we can show them ourselves using different measures or different groups of people. The fact is that many so-called findings in psychology are the results of chance, and even the statistical methods that psychologists adopt do not protect them from drawing the wrong conclusions once in a while. One alternative, therefore, is to ascribe the longitudinal change in social independence to pure chance. The finding may be meaningless, a fluke.

There is some basis for thinking that this might be so. Social independence is an aspect of introversion, and our finding might be taken as evidence that individuals become more introverted with age—an interpretation that supports some theories of aging. But the 16PF also has several scales that tap other forms of Extraversion, and none of these other scales showed a similar change in our sample. If we wanted to find evidence of change (as many developmentalists do), we might argue that independence happens to be the only aspect of Extraversion that shows maturational change. In principle there is nothing wrong with this hypothesis. If all Extraversion scales behaved identically, there would be no reason to measure them separately. But notice that in order to claim this longitudinal change as an age effect, we have to make two rather cumbersome assumptions: first, that the maturational changes do not show up in cross-sectional studies because of some mysterious confounding of generational differences with age changes; second, that there is something special about independence that distinguishes it from other closely related scales that show neither the maturational changes nor the generational differences. Possible, but unlikely.

Fortunately, there is a much more direct way to find out if this was a chance result. An independent longitudinal study carried on at Duke University over a period of 8 years also employed the 16PF, although in a slightly shorter, simplified form (Siegler et al., 1979). Researchers

looked for age changes and differences in a sample of men and women initially over 46 years old, who represented not only a different geographic region but also a somewhat broader range of socioeconomic status. These researchers also reported that their subjects improved on the second testing of intelligence, but they found neither cross-sectional differences nor longitudinal changes in social independence. Thus, the apparent decline in Extraversion was not replicated across methods (it was found in longitudinal but not cross-sectional analyses), facets (social independence but not other Extraversion scales), or samples (Boston but not Duke). Since replication is the foundation of scientific inference, the case for change in that personality trait seems closed.

The Duke investigators also reported an anomalous finding: Although there were no cross-sectional differences in Guilt-proneness, a scale that measures an aspect of Neuroticism, they found that with time men tended to score lower and women higher on this scale. None of the other Neuroticism scales showed the same pattern, nor did the men in the Boston study become less guilt-prone with age. Again, we conclude that this finding can best be ascribed to chance.

Most importantly, the Duke study confirmed the Boston study in finding overwhelming evidence of stability. It found no longitudinal changes in the degrees to which subjects were outgoing, emotionally stable, assertive, happy-go-lucky, conscientious, adventurous, tender-minded, suspicious, imaginative, shrewd, liberal thinking, independent, controlled, or tense.

SEQUENTIAL STRATEGIES: AVOIDING PRACTICE AND TIME EFFECTS

Longitudinal methods, we have pointed out, are free from some of the problems of cross-sectional studies, but they have other problems of their own. One of these we have already encountered in discussing the changes in intelligence seen in the Boston and Duke studies. When individuals take the same test repeatedly, they may answer differently simply as a result of exposure to the test. When an ability test is involved, this is called a *practice effect*, and that term has also been used to describe other effects of repeated measurement. For example, people may find a personality test less threatening the second time and give more candid answers. Or they may find it boring and become careless.

These differences might show up as “longitudinal changes” when the results of the first and second tests are compared, even though they have nothing to do with the aging process.

On the other hand, some critics have pointed out that individuals often like to give an impression of consistency. They may recall what they said the first time when asked to respond again, and they may try to give the same answers. This would give the impression of stability even in the face of real change.

Test makers are painfully aware of this problem and have come up with a few partial solutions. One is to use a different but similar test the second time—a so-called *parallel form*. This is the solution adopted, for example, by the Educational Testing Service when students take the Scholastic Aptitude Test (SAT) on more than one occasion. Since each test is different in content, there is no direct possibility of remembering the questions and figuring out the answers between tests. But most students will score a little higher the second time they take the test, probably because they feel more comfortable with the process. Parallel form testing is not a perfect solution.

If we are really concerned with the growth and change of a particular individual, the problem of practice effects is unavoidable. But if we are only interested in finding out whether people change with age, an alternative strategy has been advocated by a number of prominent developmental methodologists (P. B. Baltes, Reese, & Nesselroade, 1977; Schaie, 1977). Suppose we take a random sample of individuals born in 1920 and randomly divide them into two groups. We know from the laws of statistics that, if we have a large enough sample, the two halves are unlikely to be different on any variable we choose to measure; they are equally representative of the population of men and women born in 1920. Next, suppose we measure the personality of the first group in 1980 but do nothing to the second group. If we return in 2000 to measure the first group, we will have the standard longitudinal design and we will never quite know if the differences we find (or fail to find) are the result of aging or of practice effects. However, if in 2000 we measure the *second* group and compare it with the first group as measured in 1980, the difference cannot be the result of practice (because the second group had not taken the test before) and may therefore be a consequence of aging itself. (Incidentally, if we measure both groups in 2000, we can compare changes for the groups with and without practice and so estimate what the pure practice effects are. That is extremely useful information for other investigators who need to know whether

they should worry about practice effects or whether they can safely ignore them.)

This design, in which different individuals born at the same time are measured at two or more different times, is known as a “cross-sequential design with independent samples,” or, as we will call it, a *cross-sequential* design. Douglas and Arenberg (1978) reported cross-sequential analyses on the GZTS scales. Their two samples were collected at successive intervals (1958–1968 and 1968–1974), making this kind of analysis possible. Their second sample scored lower than their first on Emotional Stability, Personal Relations, and Masculinity, suggesting an increase in Neuroticism. Note that this finding appears to contradict the small cross-sectional decline in Neuroticism shown in Figure 2. None of the other GZTS scales showed significant cross-sequential effects.

A related approach is called a *time-sequential design*: Here individuals of the same age are measured at different times: 60-year-olds, for example, might be surveyed in 1980 and in 2000. The differences could not be caused by aging itself, since both groups are the same age. They could, however, be the result of generational differences, since the 60-year-olds of 1980 were born in 1920 whereas the 60-year-olds of 2000 were born in 1940. A great deal of ingenuity went into the invention of these designs, and they elegantly solve some of the vexing problems of aging research. But they too are ambiguous, for a reason that has not yet been mentioned.

Take as an example attitudes toward women. Suppose we had asked a national sample of 20-year-olds in 1960 if they believed that women should have their own career, freedom of choice in family planning, or the right to participate in professional sports. Relatively few would have endorsed these aspirations. But if we had asked the same questions of a different group from the same cohort (now 60-year-olds) in 2000, presumably a much larger proportion would have agreed (Costa & McCrae, 1982). We know this is not a cohort effect, or generational difference, since all these individuals were born in the same year. We know it is not a practice effect, since the respondents in our second sample have never been interviewed before. But surely we would not conclude that it was an aging effect, that individuals adopt increasingly liberal attitudes toward women as they grow older. It is more likely that we would insist the *times* had changed and almost everyone had modified opinions on these issues as a result of the Women's Movement. This would be a clear example of what is known as a *time-of-measurement effect*.

Pollsters are as plagued by time-of-measurement problems as educators are by practice effects. Some theorists (P. B. Baltes & Nesselroade, 1972) have claimed that these problems are less crucial to personality investigators because personality is unlikely to be shifted by temporary historical movements or swings in attitude. Perhaps so, perhaps not. In view of how little we know about adult personality, many would argue that it is not safe to make that assumption. What we need is a design that avoids the problems of the cross-sequential design—one that can compare people at different ages but in which we can rule out the possibility that time of measurement is responsible for the apparent effect. Is there any way to do this?

There is one way. We can eliminate time-of-measurement effects (and practice effects) if we measure all our subjects only once, all at one time. Of course, in order to see the effects of age, we would have to measure simultaneously individuals of different age groups. And this, lo and behold, is nothing other than the cross-sectional design with which we started. We have come full circle.

This curious circumambulation has bedeviled researchers for years and left the field in some perplexity. A number of additional designs of greater complexity have been proposed to break out of the circle, but critics have always been able to weave them back in. The fact is that individuals are always born in one and only one generation, and they always grow older during a certain historical period when things are changing in a certain way and to a certain degree. Short of cloning people to raise them in different circumstances, we find no way out of the dilemma. Time, cohort, and aging are inextricably confounded.

AN INTEGRATED APPROACH

We might at this point abandon all hope and move to another field of study. Alternatively, we can adopt a more constructive point of view (Costa & McCrae, 1982). Although there is no way to be sure that we interpret an effect correctly, there are ways to improve our chances. For example, when we see the same direction and magnitude of an effect in cross-sectional, cross-sequential, and longitudinal studies, it becomes highly likely that a real aging effect has been found. Exactly this pattern was found when analyses of health problems were conducted (Costa & McCrae, 1980b). All types of analyses pointed to an increased frequency

in men of sensory, genitourinary, and cardiovascular complaints with age. These results are hardly surprising, but they do demonstrate that the analyses work when a true maturational effect is there to be found.

The Douglas and Arenberg (1978) study of GZTS scales and a more recent study of NEO-PI scales provide examples of the kinds of inference that can be made when multiple analyses are conducted in the same sample.

Longitudinal analyses of the GZTS showed that 5 of the 10 scales declined between the first administration and the second, some 7 years later. But the results of other analyses indicated that only two of these were likely to be aging effects. Thoughtfulness and Personal Relations shared a pattern of effects in longitudinal, cross-sequential, and time-sequential designs but not in the cross-sectional analysis. The longitudinal effect may have been the result of aging, time, or practice; the cross-sequential effect, of aging or time; and the time-sequential effect, of cohort or time. The one common element in the three designs is time of measurement, so it seems likely that during the decade of the 1960s participants become less thoughtful and less trusting of others. This line of reasoning is confirmed by the failure to find any cross-sectional differences in Thoughtfulness or Personal Relations: When everyone is measured at the same time, there are no age or cohort differences.

Friendliness shows a more complex pattern. Like Thoughtfulness and Personal Relations, it declined in the three designs that share time of measurement as an effect, but it also showed a cross-sectional *increase*. Douglas and Arenberg (1978) interpreted this as evidence not of maturation but of long-term cultural change: "Successive cohorts and individuals became more easily aroused to hostility and tended to be less agreeable" (p. 745). Certainly we would want to see replications of this finding before accepting it as valid.

Only two traits survived this screening: Masculinity and General Activity level, both of which declined with age in cross-sectional, longitudinal, and cross-sequential designs but showed no effect in the time-sequential analysis. Results were just what one would expect if there were maturational changes in Masculinity and Activity, and the only alternative explanations involved complicated and improbable combinations of practice, time-of-measurement, and cohort effects.

However, the changes in both General Activity level and Masculinity on the GZTS are extremely small. For the statistician, the statement that they amount to about one-eighth of a standard deviation over 7

years communicates that point. For others, another comparison may make it more forcefully. The observed rate of decrease in Masculinity was 0.41 items every 6.6 years. We once computed that at this rate it would take our older subjects, initially 75 years old, 136 years to reach the same level of femininity as the average college woman. Psychologists may speak of the feminization of older men, but the reader should recall that the process would be fully completed only if men lived 211 years!

Even more extensive analyses were conducted on NEO-PI scales in a 6-year longitudinal study of 983 men and women ranging in age from 21 to 96 (Costa & McCrae, 1988b). In addition to self-reports, we conducted some parallel analyses on spouse ratings of 167 men and women. (Because we did not recruit a new sample of spouse raters, we were not able to conduct cross-sequential analyses of ratings, and because we had spouse ratings of A and C only at the second time period, we were unable to do longitudinal analyses on these scales.) The results of cross-sectional, longitudinal, and cross-sequential analyses are summarized in Table 8. In this table, minus signs suggest a decline with age, plus signs an increase, and zeros no effect; stronger effects are indicated by more signs. However, none of the effects accounts for as much as 10% of the variance in personality scores; age and aging are at best weakly related to personality.

Although it might appear that there are several age effects, a closer

TABLE 8. Age Effects from Analyses of NEO-PI Domain Scales

NEO-PI scale	Cross-sectional		Longitudinal		Cross-sequential
	Self-report	Spouse rating	Self-report	Spouse rating	
Neuroticism	--	---	-	+	0
Extraversion	---	---	0	0	+
Openness	---	--	0	0	+
Agreeableness	+	0	---		---
Conscientiousness	0	+++	-		-

Note. Minus signs indicate a negative association of the variable with age; plus signs indicate a positive association. Larger effects are indicated by more signs. Longitudinal analyses of Agreeableness and Conscientiousness and cross-sequential analyses were conducted only for self-reports. Adapted from Costa and McCrae (1988b).

inspection of Table 8 shows very little *consistent* evidence of maturational effect—indeed, not a single scale shows unequivocal evidence of change. Total Neuroticism, for example appears to decline in cross-sectional analyses of both self-reports and spouse ratings. It also declines over time in self-reports. But according to husbands and wives, rated Neuroticism actually increased over the 6-year period! No effect was seen in the cross-sequential analyses.

All the other studies we have discussed—and most in the literature—are based on self-reports, and systematic biases in self-reports might account for the general lack of age changes reported. It could be argued that external observers may more accurately perceive and report true aging effects. The analyses of spouse ratings summarized in Table 8, however, confirm the absence of strong and clear maturational effects for the three domains of Neuroticism, Extraversion, and Openness.

For a look at all five factors, consider a 7-year study of peer ratings (Costa & McCrae, 1992b). In 1983 we had asked BLSA participants to nominate friends or neighbors who knew them well, and then asked these peers to provide personality ratings of the BLSA participants. In 1990 we recontacted peer raters; we also recruited a new group of raters for new BLSA participants (recall that an independent sample is needed for cross-sequential analyses). Cross-sectional analyses showed very small effects: N, E, and O declined with age; A and C increased (although the effect for C was not significant). Longitudinal analyses showed no changes for men but a small increase in Conscientiousness for women. Finally, cross-sequential analyses showed no significant effects for any of the factors. It appears that personality as judged by friends and neighbors changes little over 7 years of adulthood.

It is worth considering what different conclusions would be drawn if the different analyses summarized in Table 8 had been done in separate studies. The researchers who conducted cross-sectional studies would have reported decreases in N, E, and O and increases in A and C; the researchers who conducted the cross-sequential analyses would have found increases in E and O and declines in A and C! Neither set of conclusions would have been replicated by the longitudinal researchers. The moral seems to be that so many factors—sampling, cohort differences, practice effects, time-of-measurement artifacts, selective mortality—affect results that most so-called aging effects are probably spurious. Only when several types of evidence are considered together can reasonable conclusions be drawn.

One consistent pattern was found when the facets of N, E, and O were separately analyzed. NEO-PI Activity showed evidence of decline in four of the five analyses. Because another scale measuring this trait also showed declines in the Douglas and Arenberg (1978) study, and because peer ratings also showed declines in Activity (Costa & McCrae, 1992b), it seems reasonable to conclude that as people age, there is a maturational decline in the pace and vigor of life. However, here too the effect is small and does not outweigh other sources of individual differences. Many 70-year-olds are more active than many 30-year-olds.

As this chapter makes clear, the simple question of "What happens with age?" is far from easy to answer. In many areas of gerontology, the separation of cohort, time-of-measurement, and practice effects from true maturational changes may require a lifetime of work. In personality, the problem is greatly simplified by the nature of the variables studied. No matter how you view it, the only consistent evidence points to predominant stability. With age, adults as a group neither increase nor decrease much in any of the traits identified by major personality instruments.

RECENT DEVELOPMENTS

Small, Slow Changes

Our integrated approach to interpreting age effects is reasonable, but it might be argued that the data we have applied it to are inadequate. The NHANES study (Costa, McCrae, et al., 1986) examined thousands of people over a wide age range. By contrast, our longitudinal and cross-sequential studies used much smaller samples and covered a much more limited portion of the lifespan; they may have lacked the statistical power to spot very small effects reliably. If there were large age effects to be discovered, our method would have found them, and we can conclude with considerable confidence that there are no large effects after age 30. But science progresses by making ever smaller and more precise measurements. If we want to find out about the possibility of small, slow changes, we need longitudinal studies that use very large samples or very long intervals. Some of both have recently been reported.

Costa, Herbst, McCrae, and Siegler (2000) conducted a longitudinal study of 2,274 University of North Carolina alumni. They were about 40 years old at first assessment and were reassessed after 6 to 9

years. Analyses showed that there were significant declines in N, E, and O, but they were tiny: about one-tenth of a standard deviation. At that rate, over the 40 years from 30 to 70, N, E, and O would decline less than half a standard deviation—an amount in line with the very small cross-sectional declines seen in Figure 2. Longitudinal declines in N and E were also reported by Loehlin and Martin (2001) in a large Australian sample using a different measure. With large samples, it is possible to see longitudinal changes that mirror the cross-sectional age differences.

The NHANES sample did not include measures of A and C, but later cross-sectional studies (Costa & McCrae, 1992a) suggested that these two increased with age. Certainly it is the case that college students are less conscientious and agreeable than adults, and so far developmental changes in later adulthood have echoed the more substantial changes in early adulthood. But the Costa, Herbst, et al. (2000) study did not support this hypothesis. Agreeableness did not change at all, and Conscientiousness actually showed a very small decline. As we show in a later chapter, there is an entirely different line of evidence that suggests that A and C really do increase across the adult years. It is fair to say that the developmental course of A and C after age 30 is currently a mystery—except that whatever changes occur are probably small and subtle.

Ravenna Helson has devoted her career to long-term longitudinal studies. Most of her samples have been relatively small, but by now she has followed some of them for as long as 40 years. In a series of analyses of CPI scales over that time interval she demonstrated significant changes consistent with decreases in N and E and increases in C (Helson & Kwan, 2000). These longitudinal changes averaged about one-third standard deviation. At least with respect to N, E, and O, there is now fairly consistent evidence of decline during adulthood.

How small a change is one-third standard deviation? One way to think of it is by a comparison with height and weight. The standard deviation of height among adults is about 3½ inches, and one-third of that is a little more than 1 inch. So if height declined at the same rate as Neuroticism, a six-foot-two 30-year-old could expect to be six-foot-one at age 70. That is a perceptible decline but hardly a dramatic one: Six-foot-one is still tall in an absolute sense. The comparison with weight is even more revealing, because most of us are acquainted with changes in weight. The standard deviation of weight among adults is about 35

pounds; imagine living for 40 years and gaining no more than 12 pounds!

Another Mystery

Recently Twenge (2000) reported a remarkable result: She collected all the studies she could find of students given standard tests of anxiety or Neuroticism between 1952 and 1993, and plotted mean values by year. Her results suggested that each successive cohort has become more anxious—by almost a full standard deviation. She suggested that the stressors of modern life have induced these generational changes in personality. Most personality psychologists are puzzled by these findings, and it may take some time to properly evaluate them.

As it happens, the earliest-born subjects in her meta-analyses correspond to the youngest respondents (age 34–54) in the NHANES sample (Costa, McCrae, et al., 1986). If each birth-year cohort is getting more and more anxious, then when they are all tested at one time as adults the oldest (i.e., the earliest born) ought to be the best adjusted. Neuroticism in Figure 2 should show a steep decline with age, not the near-flat pattern we see. Why is that decline absent? Perhaps, as in the pessimistic scenario we envisioned earlier in this chapter, it is because with age people become increasingly neurotic and the birth-cohort effects are canceled out by equal and opposite aging effects. But all the longitudinal evidence says that people actually *decline* a bit in Neuroticism with age. The cross-sequential analyses Twenge (2000) reports for her student samples are inconsistent with longitudinal and cross-sectional analyses of adults. Perhaps the stress-induced anxieties of childhood simply evaporate in adulthood. This curious set of findings awaits a final explanation.

A Different Analysis

All the analyses we have discussed so far analyzed the mean values of groups of people—different groups, or the same groups at different times. In recent years methodologists have argued that there are much more powerful ways to examine longitudinal data, especially if people are assessed not twice but on multiple occasions. Then a curve can be fit to each individual's data and these growth curves can be analyzed. For example, even when the group as a whole shows no change, it is likely

that some individuals have increased and others decreased. It would be informative to consider whether the increasers share some common experience: Perhaps they all got married or lost their jobs. Such findings might lead to a theory of the causes of personality change.

Aldwin, Spiro, Levenson, and Cupertino (2001) used such an analysis on reports of psychological symptoms (akin to Neuroticism) in the Normative Aging Study (the same sample we had studied 20 years earlier). Each of 1,515 men had been measured on an average of six occasions spanning up to 25 years. When the overall pattern was examined, “the predicted number of psychological symptoms did not change with age.” But when individual curves were examined, subgroups of men could be discerned. By far the largest group (about 75%) consisted of men who had few symptoms to begin with and stayed that way. Three smaller groups consisted of people who started with a moderate level of symptoms. Some declined a bit in midlife, then increased again; some increased until age 65, then began to decline; some peaked in the 40s, a pattern the authors likened to a midlife crisis.

These are intriguing findings, well worth replicating. One caution is needed however: The statistical techniques used piece together a curve from age 35 to age 75, a span of 40 years. Yet the average subject had been followed only for 19 years, and none had been followed for more than 25 years. Whether individuals will actually continue on their projected trajectory remains to be seen.

IMPLICATIONS: DEBUNKING SOME MYTHS OF AGING

Complexities and subtleties aside, the take-home message of this chapter is that, after age 30, the mean levels of personality traits change little. What does the evidence of stability in personality mean for our view of aging? For some reason, when we say we find no evidence of growth or decline in personality, people hear only “no growth” and regard us as the bearers of bad news. But there is good news here, too. Perhaps it is unfortunate that people do not continue to grow and develop in adulthood, but surely it is reassuring to find that they do not deteriorate. In view of popular and prevalent conceptions of aging, this is by far the more important implication.

In the past few years socially conscious individuals have attempted to create positive stereotypes of aging, reminding us that the aged are

revered in many cultures for their experience and wisdom (see the *Positive Aging Newsletter*, edited by Kenneth and Mary Gergen). Octogenarian musicians and actors and artists are showered with honors, and their current productions are uncritically acclaimed. In many ways, this newfound respect for the aged is just compensation for decades of neglect. But it often seems that there is something condescending and infantilizing about the status we grant them, as if the elderly needed special consideration. Skinner (1983) noted that the mindless veneration with which the ideas of older scholars and scientists are received reinforces platitudinous thinking and contributes to a decline in creativity.

We might even suspect that many people bend over backward to think well of the elderly because at a deeper level they fear that age is nothing but decline. There are real losses in physical strength and vigor, in sexual interest, and in certain intellectual abilities, for which it is necessary to make allowances. But there are no significant declines in personality.

We need not worry that we will become crotchety and hypochondriacal with age, or that only firm resignation can save us from despair and the fear of death. We need not anticipate increasing social isolation and emotional withdrawal from the world. There is no reason to think that our interests will atrophy or that our values and opinions will become increasingly rigid and conservative.

And just as we need not dread our own future, so we need not pity others who have already reached an advanced age. The elderly are no more emotionally vulnerable or ideologically rigid than anyone else, and giving them special treatment in these areas is unnecessary and probably unwise. Older individuals are entitled to all the respect due any human being, but genuine respect means seeing them as they really are. With regard to personality they are not much different from any other adults.

* * *

In the mid-1990s we returned to cross-sectional studies of personality that were to give us an entirely new perspective on age and personality. By that time, the FFM had become established as a model of personality trait structure not only in the United States but in many countries in Europe and several in Asia. Translations of the NEO-PI-R began to

spring up. It suddenly became possible to move beyond a parochial concern with aging in the United States to a focus on aging as a universal human process. Would the patterns of stability and change that we saw in American adults be found in South Korea, or Turkey, or Russia? Longitudinal studies in these countries would be ideal, but in the meantime the instant gratification of cross-sectional research seemed very appealing.

Cross-Cultural Perspectives on Personality and Aging

Exotic people and faraway lands have always held fear and fascination. To the imagination of the Middle Ages, the unknown regions of the world were peopled by strange monsters (see Figure 3). When, in the Age of Exploration, the inhabitants of the New World were actually encountered, they were regarded as savages, not quite fully human. In the Enlightenment, particularly in the writings of Jean-Jacques Rousseau, this evaluation was reversed and preliterate people were idealized as noble savages, living as nature intended, without the defilement of corrupt civilization. It is well to bear this bit of history in mind, because it is repeated in some respects in scientists' perceptions of personality across cultures.

Today people are interested in the manners, customs, and beliefs of other groups of people for two main reasons. First, as ever, there is an intrinsic fascination in the topic. So many ideas and values that are second nature to us are challenged by the ideas and values of other cultures; for those who are open to experience, these challenges provide a basis for rethinking our own ways. Second, progressive thought in the second half of the 20th century has increasingly emphasized the basic human equality of all peoples and affirmed the responsibility of each person to become informed about and sensitive to cultural differences. Multiculturalists (Gergen, Gulerce, Lock, & Misra, 1996) argue that broadening one's perspectives beyond the culture of one's birth is both morally and scientifically required of psychologists. If we aspire to uni-



FIGURE 3. Imagined inhabitants of the undiscovered world. From Hartmann Schedel's *Liber Chronicarum*, Nuremberg, Germany, 1493.

versal laws of personality development, it is obvious that we must test them in many different cultural contexts.

Multiculturalism grew out of a position called *cultural relativism*, which claimed that human values can only be judged from within the perspective of the culture in which they arose, because values are a cultural creation. From this perspective, we should not condemn the precolonial Aztec people of Mexico for practicing human sacrifice, because within their view of the world this was an act of piety, necessary for the propitiation of the gods and the continued survival of their society. Carried to an extreme, cultural relativism can be used to justify any action (like Pol Pot's genocide in Cambodia in the name of socialism), but the basic idea that we should not simply assume that our views and values are the right ones is still powerful.

The ideas associated with cultural relativism were popularized by anthropologists like Ruth Benedict (1934) and Margaret Mead (1928), who were among many psychologists, psychoanalysts, and sociologists concerned with the topic of culture and personality in the first half of the 20th century (McCrae, 2000b). Although there were certainly many variations (LeVine, 2001), the basic idea in this school of thought was that personality was shaped by culture, especially child-rearing practices. Because there were obvious and dramatic differences between cultures, one could expect that there would be equally great differences between the personality of, say, the Yanomamö of Venezuela and the Balinese of Indonesia. Further, because all Yanomamö shared a common culture, they should also share a common personality, sometimes called the *modal personality structure*.

Personality assessment was in its infancy then, and most of the em-

pirical tests of this hypothesis used projective methods like the Rorschach Inkblot Test. As we show in Chapter 8, there are good reasons to be skeptical about Rorschach results, even when the test was administered in the culture in which it was invented. Applied abroad, the results were almost certainly meaningless. They did, however, point to one conclusion: People in the same culture do not share a single personality style; there are individual differences everywhere.

For a number of reasons, including the failure of the modal personality hypothesis, both anthropologists and psychologists drifted away from this topic by the 1960s. One legacy of this period, however, was a widespread belief among social scientists that personality must be fundamentally different in markedly different cultures. As recently as 1996 Samuel Juni expressed this view unequivocally:

The simplistic (a posteriori) basis of the Five-Factor Model, as it is derived from colloquial usage of language, makes the model and its tools intrinsically bound to the culture and language that spawned it. Different cultures and different languages should give rise to other models that have little chance of being five in number nor of having any of the factors resemble those derived from the linguistic/social network of middle-class Americans. (p. 864)

Just as Medieval Europeans imagined the undiscovered peoples of the world would be monsters, so many contemporary psychologists believed that personality structure in foreign cultures would be entirely different from the familiar FFM.

In the absence of data, that is perhaps a reasonable presumption. But it would be unfortunate if it were true, because cross-cultural comparisons are potentially invaluable as tools for understanding the FFM. We cannot normally conduct experiments to discover the determinants of personality traits; it would be both unethical and unfeasible to manipulate styles of child rearing or levels of prosperity or religious disciplines to see if they affected personality structure. But we can compare individuals in cultures that differ in these regards, and we can potentially learn a good deal from these “natural experiments.”

A Universal Structure

But it would only be possible to compare cultures if the same factors are found in both. Does lenient discipline in childhood lead to greater

Openness to Experience in adults? It would make no sense to ask that question in a culture in which there was no Openness factor. Consequently, the first question that must be asked in cross-cultural comparisons is whether the structure of personality is basically the same.

There are even more basic questions than this: We could ask whether the specific traits that define the factors are universal. Perhaps there is no Openness factor in Culture X because there are no consistent individual differences in intellectual curiosity or aesthetic sensitivity. There is a yet more basic question: "Do people from all cultures have personality traits?" It has been argued (see Church & Katigbak, 2000) that traits are Western inventions, appropriate only to highly individualistic people who define themselves by their personal attributes. In collectivistic cultures like Japan or Zimbabwe, people may define themselves in terms of their interpersonal relationships. Perhaps for such people the concept of *trait* is irrelevant.

Although these may seem to be abstract and difficult questions, there is a straightforward way of addressing them. If we translate the scales of the NEO-PI-R into different languages and administer them to samples from other cultures, we can try to replicate the factor structure. If the whole notion of traits is meaningless in a given culture, we should find no structure at all. If, on the other hand, we find a structure resembling the FFM, it would mean not only that personality trait structure is generalizable but also that the individual traits are similar in different cultures. If Openness to Ideas loads on the Openness factor in culture X, then there must be systematic individual differences in intellectual curiosity in culture X. During the past decade dozens of studies have addressed these issues, and the answer is now clear.

As long ago as 1983, Sybil Eysenck showed that two factors, Neuroticism and Extraversion, could be found in many cultures. Lexical researchers in Germany and the Netherlands found versions of the FFM in their analyses of Dutch- and German-language trait descriptive adjectives (Brokken, 1978; Ostendorf, 1990). But German and Dutch are closely related to English, since all are from the Germanic branch of the Indo-European language family. Perhaps the FFM is distinctive to Germanic languages.

The universality of the FFM was first clearly established when the same factor structure was found in the NEO-PI-R when translated into languages from several different language families—Indo-European, Hamito-Semitic, Sino-Tibetan—as well as Japanese and Korean, lan-

guages that are not classified into any larger language family (McCrae & Costa, 1997b). Since that time, the FFM has been replicated in languages from the Bantu, Malayo-Polynesian, Uralic, and Altaic language families, in more than two dozen different cultures (McCrae, 2000a). Table 9 illustrates the findings. In it, the adult American factor structure (Costa & McCrae, 1992b) is reproduced alongside the factor structure found in Indonesia (Halim, 2001) in a small sample of women college students. Although the two structures are not identical—for example, Excitement Seeking is more strongly related to Openness than to Extraversion in the Indonesian sample—it is clear that the same five factors are found in both samples, despite differences in language, culture, and age of the samples.

Curiously, this discovery has been greeted with mixed feelings by the community of cross-cultural psychologists. It is considered by some a form of Western imperialism that an instrument developed by Americans in English to describe the personality structure of Americans has been imposed on the rest of the world. The imputation of conquest is seen in Michael Bond's (2000) whimsical assertion that "with the publication of McCrae and Costa's (1997b) article in the *American Psychologist* titled 'Personality Trait Structure as a Human Universal,' the pacification of the known personological world may now be complete" (p. 63). That sentiment aside, critics make the legitimate point that this research does not address the possibility that there are other, indigenous, factors that may be unique to other cultures. For example, the NEO-PI-R does not include scales to measure Chinese Tradition (Cheung & Leung, 1998), so there is no possibility that a Chinese Tradition factor would emerge, even when the instrument is administered in Chinese. A conservative interpretation of the data to date would be that personality traits in all cultures studied so far include the FFM, although some may also have other dimensions of personality.

Western civilization does have a history of imperialism that was often brutal to indigenous cultures, and cross-cultural psychologists are right to be vigilant about repeating those misdeeds in science. But Western intellectuals also have a tradition of idealizing non-Western cultures, from the notion of the noble savage to present-day claims of the irreducible uniqueness of different cultures (Shweder & Sullivan, 1990). In the long run, it may be that the best way to promote human understanding is not only by celebrating diversity but also by showing that all

TABLE 9. Factor Structure of the NEO-PI-R in American and Indonesian Samples

NEO-PI-R facet	N		E		O		A		C	
	U.S.	Indo.	U.S.	Indo.	U.S.	Indo.	U.S.	Indo.	U.S.	Indo.
N1: Anxiety	81	74	02	-19	-01	-20	-01	17	-10	-13
N2: Angry Hostility	63	61	-03	04	01	-06	-48	-48	-08	-15
N3: Depression	80	74	-10	-26	02	02	-03	04	-26	-17
N4: Self-Consciousness	73	60	-18	-34	-09	-26	04	-00	-16	-08
N5: Impulsiveness	49	54	35	37	02	-10	-21	-18	-32	-34
N6: Vulnerability	70	69	-15	-04	-09	-12	04	00	-38	-37
E1: Warmth	-12	-16	66	69	18	19	38	17	13	10
E2: Gregariousness	-18	-17	66	66	04	01	07	08	-03	-06
E3: Assertiveness	-32	-39	44	51	23	15	-32	-36	32	24
E4: Activity	04	-17	54	57	16	17	-27	-31	42	18
E5: Excitement Seeking	00	-13	58	30	11	46	-38	-37	-06	-05
E6: Positive Emotions	-04	-14	74	62	19	39	10	-05	10	08
O1: Fantasy	18	-03	18	20	58	57	-14	-05	-31	-01
O2: Aesthetics	14	19	04	11	73	67	17	07	14	04
O3: Feelings	37	43	41	28	50	48	-01	00	12	17
O4: Actions	-19	-48	22	33	57	27	04	-15	-04	-11
O5: Ideas	-15	-12	-01	21	75	63	-09	-15	16	21
O6: Values	-13	-21	08	-16	49	63	-07	07	-15	07
A1: Trust	-35	-30	22	24	15	04	56	61	03	-03
A2: Straightforwardness	-03	-04	-15	-18	-11	-09	68	65	24	13
A3: Altruism	-06	-15	52	39	-05	13	55	54	27	29
A4: Compliance	-16	-15	-08	-27	00	-09	77	63	01	-12
A5: Modesty	19	17	-12	-40	-18	-05	59	41	-08	-14
A6: Tender-mindedness	04	18	27	47	13	-12	62	44	00	-02
C1: Competence	-41	-25	17	19	13	04	03	10	64	69
C2: Order	-04	02	06	-08	-19	-11	01	-09	70	74
C3: Dutifulness	-20	-22	-04	-08	01	13	29	23	68	68
C4: Achievement Striving	-09	-15	23	19	15	05	-13	-08	74	72
C5: Self-Discipline	-33	-26	17	05	-08	04	06	02	75	71
C6: Deliberation	-23	-21	-28	-31	-04	-08	22	20	57	54

Note. Principal component loadings; decimal points omitted; loadings .40 in absolute magnitude in bold-face. U.S. = American normative data, $N = 1,000$; adapted from Costa and McCrae, 1992b; Indo. = Indonesian data from a Procrustes rotation of Halim (2001, Table IV.2), $N = 160$.

human beings are fundamentally alike. The universality of the FFM is an important contribution to that effort.

ADULT DEVELOPMENT ACROSS CULTURES

In the mid-1990s colleagues in Europe developed a new translation of the NEO-PI-R. They followed our standard procedure: Psychologists native to the culture translated the items, with more attention to conveying the constructs than to producing a literal, word-for-word translation. Other scholars, unfamiliar with the NEO-PI-R, then translated the new version back into English. We reviewed the English back-translation to make sure the sense had been preserved. When there appeared to be problems (typically in less than 10% of the items), we asked them to try again, and the process was repeated until an acceptable back-translation had been achieved.

These colleagues then administered their instrument to a sample of students and their parents, and sent us a table of means and standard deviations for the two samples. At the time, we had no experience comparing average or mean levels across cultures, so we could not comment on what the mean levels meant. But we did notice that there was something odd about the data. When the students were compared to the adults, they scored lower on Neuroticism, Extraversion, and Openness, and higher on Agreeableness and Conscientiousness—students appeared to be more mature, serious, conservative, helpful and disciplined than their parents. This is, of course, the opposite pattern from that found in Americans.

For all we knew at that time, personality development might be very different in the United States and in Europe. Still, it seemed odd that the pattern was exactly reversed, so we considered another hypothesis. Was it possible, we asked our colleagues, that they had scored the test wrong, such that high and low scores were reversed? A week later we got a reply: They had checked their scoring program, and, yes, it had been reversed. New, corrected mean levels showed that students were actually higher in Neuroticism, Extraversion, and Openness, and lower in Agreeableness and Conscientiousness.

A few days later the full import of that news dawned on us: When scored correctly, this European sample showed *exactly the same* pattern of personality development we had seen in Americans. Was this just a

coincidence, or did it mean that personality development was a universal pattern?

We immediately turned to the psychological literature to see if we could confirm this hypothesis. Ideally, we would have liked to find longitudinal studies tracing personality traits from adolescence into adulthood. But we soon discovered that there was no literature to consult. There had only been a handful of longitudinal studies conducted outside the United States (e.g., Thomae, 1976); few of them had used standardized personality tests, and none of them had used measures of the FFM. More surprising was that there were also few cross-sectional studies. Although personality scales have been used around the world at least since the 1970s (Cattell, 1973), it appeared that no one had thought to use them to assess adult age differences.

Fortunately for us, we had colleagues in other countries who had also collected data from students and adults. In our first article on the topic, we analyzed data from South Korea, Italy, Germany, Croatia, and Portugal (McCrae et al., 1999). In each culture, we divided the respondents into four groups: age 18–21, 22–29, 30–49, and 50+. Figure 4

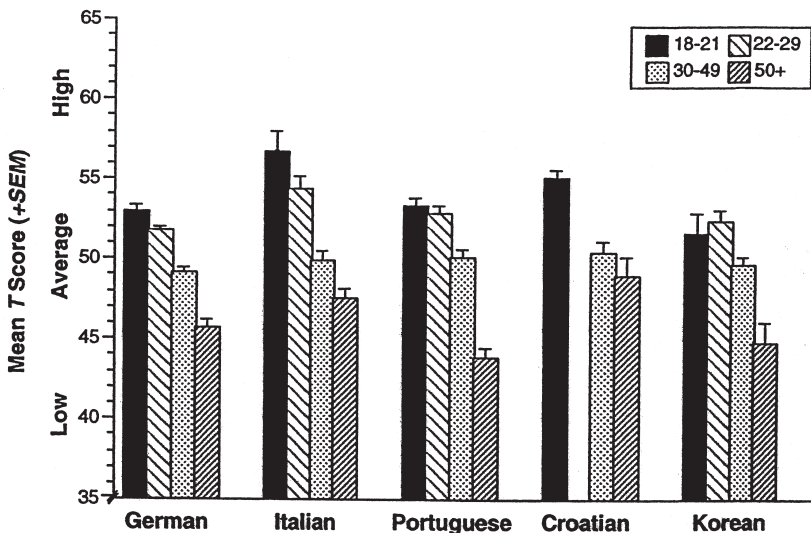


FIGURE 4. Openness to Experience for four age groups in five cultures; none of the Croatian respondents was aged 22–29. Adapted from McCrae et al. (1999).

shows how the four groups differed in Openness. In each culture, Openness was highest in the youngest groups and lowest in the older groups. These age differences are significant in all five cultural groups. In the case of Neuroticism, there were no significant age differences in the Italian and Croatian groups, but the German, Portuguese, and South Korean samples followed the American pattern. All five cultures replicated American findings for Extraversion, Agreeableness, and Conscientiousness.

History, Culture, and Cross-Sectional Comparisons

These are, of course, cross-sectional comparisons, and (as we discussed in the last chapter), there are good reasons to be cautious in inferring maturational changes from cross-sectional age differences. In particular, there is always the possibility that differences are cohort effects. The older respondents in Figure 4 may always have been lower in Extraversion and higher in Conscientiousness because they were born in an earlier era when these traits were shaped. Ordinarily, this argument cannot be challenged except by longitudinal data, and it will be years before extensive longitudinal data from other cultures are available.

But cross-cultural research is not ordinary research. Within any one culture, age and birth cohort are hopelessly confounded, but this need not be the case when data from two or more cultures are compared. As Riley, Johnson, and Foner (1972) pointed out long ago, in interpreting cohort effects, "our concern is not with dates themselves, but with the particular sociocultural and environmental events, conditions, and changes to which the individual is exposed at particular periods" (pp. 419–420). People growing up in South Korea were exposed to very different events, conditions, and changes than people growing up in the United States or Portugal (see Table 10 for a synopsis of history in these five cultures), so one might expect a very different pattern of cohort effects. If life history shapes personality traits, and if histories differ, then cross-sectional age comparisons should also differ.

But as Figure 4 shows, that is not the case. The same general trends are found in each of these five cultures. Later studies showed that the pattern continues in Russia, Estonia, Japan, the Czech Republic, Great Britain, Spain, Turkey, and China (Costa, McCrae, et al., 2000; McCrae et al., 2000; Yang, McCrae, & Costa, 1998). Consider how different life experience has been for people growing up in these countries. After

TABLE 10. An Overview of Recent History in Five Cultures.

Country	1915–1945	1945–1965	1965–1995
Germany	Defeat in World War I; Great Depression; decline of democracy and rise of Third Reich (Hitler); territorial expansion; the Holocaust; defeat in World War II	Marshall Plan and postwar reconstruction; Soviet blockade of Berlin and division of Germany; Cold War, nuclear threat, Berlin Wall; economic prosperity; television	Student movement and liberalization of values; peace and prosperity; East/West rapprochement and reunification. Current concerns: unemployment, AIDS, environmental pollution
Italy	Social and labor unrest, postwar nationalistic discontent leads to rise of fascism (Mussolini); conquest of Ethiopia and occupation of Albania; alliance with Germany brings Italy into World War II; defeat and civil war of liberation	Marshall Plan and postwar reconstruction; industrialization and urbanization foster rapid transformation of society and economic growth	Center–left coalition; social conflict, student movement, and strong unionization; legalization of divorce; escalation of political conflict; terrorism; fighting mafia and political corruption; emergence of new parties
Portugal	Brief democracy; World War I; 1926 military coup and dictatorship (Salazar); restrictions on free speech and social development; neutrality but hardships in World War II	Initial industrialization in 1950s; beginning of colonial wars in Africa, 1961, and return of colonists; continued dictatorship; widespread emigration to northern Europe	Increasing political opening; end of colonial war and massive immigration of ex-colonists; student movement in 1960s; restoration of democracy, 1974; enters European Economic Community (EEC); social liberalization; divorce legalized; women's rights upheld
Croatia	Austro-Hungarian Empire dismantled; multiethnic Yugoslavia formed under Serbian king; Croatian independence movement; occupation by Axis powers and creation of independent Croatia under fascist puppet dictator in World War II	Nonaligned Communist government (Tito) reunites Yugoslavia; gradual economic recovery; continued nationalist sentiment	Economic boom, then decline in 1980s; democracy restored in Yugoslavia, 1990; war of Croatian independence, recognized in 1992; continuing Balkan conflict

(continued on next page)

TABLE 10 (*continued*)

Country	1915–1945	1945–1965	1965–1995
South Korea	Japanese occupation; social discrimination against Koreans and Korean culture, severe economic hardship, widespread starvation, continued through World War II	Intensive efforts at economic recovery; Korean War and the partition of Korea; military dictatorship and social repression	“Economic miracle” introduces prosperity; continued military threat from North Korea; continuing dictatorship sporadically resisted by demonstrations; civilian democracy established in 1990s

Note. From Costa and McCrae (2000b).

World War II Japan and South Korea went through “economic miracles” for several decades. In that same period, Czechoslovakia and the Soviet Union underwent a period of economic stagnation and decline. Life in the 1960s was relatively uneventful for Germans compared to the upheavals of the Cultural Revolution in Mao’s China. Yet in all these countries, by the mid-1990s, adolescents differed from adults in the same ways.

Some Possible Interpretations

What are we to make of this cross-cultural uniformity in age differences? There seem to be three plausible explanations. The first is that these are cohort differences attributable to historical events that all these countries shared in common. For people all over the world, the second half of the 20th century was better than the first half, filled as it was with World Wars I and II and the Great Depression. Health and nutrition have improved for almost all countries. Television was virtually unknown to the public in the 1930; by the 1970s it was ubiquitous. Western influence and culture have permeated most of the world, with a concomitant decline in local traditions. Perhaps these influences have produced generational effects on personality traits that are common to all the cultures studied.

Although that possibility cannot be ruled out, on careful thought it is not very persuasive. It is hard to think of a good reason why television viewing would make more recent generations more anxious and

depressed than their parents, or why better diet would lead to lower Agreeableness and Conscientiousness. It is also hard to explain why events that many countries share have major effects on personality traits whereas events that are unique to individual countries seem to have so little influence. Was the introduction of Coca Cola to the People's Republic of China really more important than the Cultural Revolution in shaping generations' personalities?

A second, and more promising, explanation is that cultures everywhere develop similar social structures that in turn shape personality. The cross-sectional age differences we see in Figure 4 would then reflect real changes, but changes that are dictated by social causes. Adolescents in every culture are expected to take adult responsibilities as they grow older, including work and raising children. It is reasonable to argue that people may respond to these social pressures by becoming more serious, task oriented, and prosocial. There are several ways to test this hypothesis. For example, the age at which children are supposed to begin a family varies considerably across cultures. Does maturation in personality occur earlier in those cultures in which families are begun earlier? Within a single culture, one could conduct before-and-after studies on adolescents who do and do not begin a family (Neyer & Asendorpf, 2001).

There is, however, another possibility. Perhaps cross-sectional differences reflect intrinsic maturational changes—changes built into the human species in the same way that puberty and menopause are. In Chapter 10 we discuss in more detail the argument that age-related changes in the mean levels of personality traits are genetically determined, but it is worthwhile to mention it here. One of the major themes of this book is that personality is largely stable in adulthood, despite the many significant events that occur. Our findings on cross-cultural age differences make a related point: Maturation from adolescence to adulthood occurs in a predictable pattern regardless of the cultural context in which it occurs. In both cases, personality traits seem to follow their own path, without regard to pressures of the environment. This suggests that they are controlled by other forces, and genes might be responsible.

Why would people have genetic programs that cause them to become more introverted and less antagonistic as they enter adulthood? We may never know. It is clear that some men have genes that make them go bald and most have genes that turn hair gray, but it is hard to

explain why. It is useful to recall that evolution is driven both by random mutations and by the adaptive consequences of the mutation's expression. If there are no consequences—if going bald doesn't decrease one's chance of surviving and reproducing—then the mutation will be passed on, even though it serves no useful function.

But we can at least speculate that there might have been an evolutionary advantage in personality trait maturation. The main adaptive tasks faced by adolescents are becoming independent adults, finding a mate, and making a place in the world. Openness to Experience and Extraversion both make individuals more adventurous and help them explore the natural and interpersonal worlds; they may thus be helpful in solving the tasks of adolescents. Once settled in, however, with a home, a mate, and a family to raise, these traits become less valuable—in fact, they may be distractions. Instead, love and work should now become priorities, and these are related to Agreeableness and Conscientiousness. It is less clear why adolescents should be high in Neuroticism compared to adults, but it is hard to complain about its decline!

Cross-Cultural Evidence on Stability

Although we have noted from the beginning of this book that there is change in personality traits during some portions of the lifespan, the emphasis has been on evidence of stability after age 30. In this chapter we have argued that age differences seen in cross-cultural studies provide another line of evidence that there are intrinsic maturational patterns in the mean levels of personality traits: From adolescence to full adulthood, men and women decline in Neuroticism, Extraversion, and Openness, and increase in Agreeableness and Conscientiousness. But do cross-cultural studies also support the idea that changes after age 30 are relatively small?

In general, yes. Averaged across the five factors and the five cultures, the mean rate of change was about one-sixth standard deviation per decade, and much of this change occurred in the first decade, between 18 and 30. However, there are some notable exceptions. In Figure 4, the Italian and Croatian data follow the American pattern closely, but there appear to be substantial drops in Openness between the two oldest groups in the Portuguese and Korean data. Here a cohort effect is plausible, because in both Portugal and South Korea there are very marked cohort differences in education. The oldest subjects in both

these samples had virtually no formal education. When years of education was statistically controlled in the Portuguese sample, the difference between the two oldest samples was substantially reduced. (Education was not assessed in the Korean sample.)

The very modest relation of age to trait levels was confirmed in a later study of German, British, Spanish, Czech, and Turkish samples, where the median correlations with age ranged from $-.08$ for Openness to $.23$ for Conscientiousness (McCrae et al., 2000).

Longitudinal studies of FFM traits are currently in progress in Germany and Russia, and include both self-reports and peer ratings. Those studies should help resolve any remaining ambiguities in interpreting possible cohort effects (like the effect of education on Openness in Portugal). What is now unambiguously clear is that the traits of the FFM are found throughout the world and show much the same pattern of age differences and stability from adolescence through adulthood. Whether these patterns reflect universal social structures or characteristics of the human species, they emphasize an important way in which all people are alike. However different our cultures and historical circumstances, our dispositions follow the same course.

* * *

In Chapter 4 we contrasted longitudinal with cross-sectional studies by pointing out that the former traces development over time in the same individuals. We have now reported cross-sectional results from around the world, but we have not yet presented data from a single individual, only from groups of individuals averaged at two points in time. Surely out of a large group of people there must be some who change: a few who learn and grow from their experiences, a few who fall into despair or stagnation. Age does not bring much by way of universal decline or growth in personality traits, but perhaps it brings idiosyncratic changes that mirror the unique life courses of aging men and women. Longitudinal studies can address such questions, but only with a different kind of analysis. We turn to these issues next.

The Course of Personality Development in the Individual

We have used the word *stability* repeatedly with the assumption that its meaning is obvious. But in fact stability has several meanings, and a discussion of them is a necessary prologue to a fuller look at the data on personality and aging. Different kinds of statistical tests and sometimes different methods of collecting data are necessary when one is looking for different types of stability and change. Some kinds of stability are frankly uninteresting; others form the basis for a whole new way of thinking about the course of human lives.

Two Different Questions: Stability and Change in Groups and in Individuals

In Chapter 4 we reviewed at some length the evidence of stability of mean levels of personality traits. In the simplest case, the cross-sectional study, we found that old people generally scored neither much higher nor much lower than young adults on a variety of personality measures. This failure to find marked change in personality contrasts with the evidence of age-related changes in a number of functions. In childhood, for example, intelligence, vocabulary, and physical size and strength obviously increase through adolescence. In later adulthood,

declines in strength, memory, and hearing are equally well documented. Most personality traits, on the other hand, do not show a marked pattern of rising or falling as people age.

In all these examples, whether of change or stability, we are comparing the average levels of groups of individuals. There are good reasons for concentrating on the average (or mean) level of a trait when we are interested in the effects of age. We know that most traits show a wide distribution, with some individuals high, some low, and many intermediate in the degree to which they manifest the trait. We can rarely make assertions of the form "All 80-year olds are higher in *X* than all 70-year-olds." If a psychologist is careless enough to say that old people have poor memories, he or she is sure to hear dozens of stories about old people with good memories or young people with worse ones. When so pressed, the psychologist is likely to rephrase the statement to say that, *on average*, old people have poorer memories. "Of course," he or she might continue, "some people have better memories, some worse; they may have been born that way, or they may have developed skills through practice or education. Medications or illnesses or psychological states may interfere with memory performance. But I'm not interested in any of those things. I'm interested in the effects of age on memory. So my strategy is to measure a large group of people, some old, some young. Some of the old ones will be naturally bright, some naturally not so bright; some will be well-educated, some poorly educated. When I look at the group average, all those differences will cancel each other out. The only thing all these people have in common, and thus the only thing that will be characteristic of the group average, is the fact that they are older. The same logic applies to the young group. When I compare the average older subject with the average younger subject, any differences must be due to age."

We have already challenged the assumption that the only respect in which two such groups differed systematically was age, pointing out that they might well also differ in average education, health status, or other features. But if we can assume that those other characteristics have been controlled through a careful selection of subjects, the logic of cross-sectional comparisons is sound. Individual differences in memory are attributable either to age or to some other factors. The interest in studies of this sort is always in age as a source of individual differences; other differences are something of a nuisance, a source of possible confusion. In statistical models, they contribute to what is called an *error*

term. In fact, if everyone started out with the exactly the same capacity for memory, it would be much easier to see the effects of age.

But people do not start off with the same abilities, just as they do not start off with the same levels of Anxiety or Assertiveness or Openness to Ideas. In early adulthood we find wide variation in all the personality traits of interest to us. We find the same range of differences among old people. We know from the studies in Chapters 4 and 5 that people on the average do not change much during adulthood; but so far we have said very little about what individuals do. Ms. Smith may be docile and traditional as a newlywed, but 30 years later she may have become assertive and unconventional. Mr. Jones may be interested in automotive mechanics and sports during his 20s, but he may be devoted to Bible reading by the time he is 60. Then again, Ms. Smith may still be docile and traditional and Mr. Jones still interested in sports when both are past 70.

This is a question of stability or change not of a group of people but of individuals. The study of changes in individuals is considerably more complicated and correspondingly more interesting than the study of changes in groups. If the group as a whole changes, we can be sure that at least some of the individuals have changed, but the converse is not necessarily true. The average level of the group may not be altered even though all the individuals change—if, for example, there are as many who increase as decrease. Finding out that groups are stable, then, does not rule out the possibility that individuals change; and, in fact, a number of fascinating possibilities for individual development are consistent with the findings of group stability.

There is a major methodological difference between the study of groups and the study of individuals. To make inferences about mean changes with age, we need only compare groups of different ages (provided we can be relatively sure they do not differ in any other characteristics). Any two groups of people will do, so the cross-sectional design is a convenient, if not foolproof, method of obtaining a quick answer. We need not wait for the passage of time.

But a study of individual changes requires that we examine the same person at two or more ages: Longitudinal studies with repeated measurements of the same subjects are essential. As a result, the evidence available on which to base conclusions is much slimmer and of much more recent vintage. Only in the past two decades have more than a handful of studies been published in which longitudinal data on

adult personality were reported. Fortunately, all these studies have been in substantial agreement, and so we can draw our conclusions with considerable confidence.

There is, however, one approach in this area that parallels the cross-sectional study as a quick method of seeing what changes occur to an individual over the lifespan. In the retrospective study, people are asked to recall what they were like at earlier ages and compare their present state. On the whole, psychologists have been extremely leery of this kind of research (e.g., Halverson, 1988). Memory, they maintain, has ways of playing tricks on people. In fact, a number of studies have shown that people can and do distort the past, whether consciously or not. Many more people, for example, “remember” voting for the winning candidate than election results could possibly allow. There has been very little research comparing memories of earlier personality with actual records. However, the recent harvest of longitudinal findings allows us to compare retrospective reports with objective facts. As we will see, they suggest that conclusions based on retrospective reports generally square with prospective longitudinal findings.

DEVELOPMENTAL PATTERNS IN INDIVIDUALS

Although most personality theorists assume that adult personality is in some way an outgrowth or developmental continuation of earlier personality, other views are also possible. We have already encountered (and, we hope, countered) the critical view that wholly denies the reality of personality, assigning to it the status of a social myth or a metaphysical existence like that of the soul. So elusive and insubstantial an entity could hardly be said either to change or to stay the same, and the whole issue would be banished from the realm of scientific inquiry. A more moderate position might hold that personality is like mood—a real enough phenomenon, but one that comes and goes according to laws so obscure that it seems completely random. Finally, a number of contemporary psychologists would probably conceptualize personality as largely a function of the recent environment. They might see Extraversion, for example, as a set of learned responses to social situations: One’s environment may change radically as one goes from one’s parents’ home to college to one work situation and then another and finally to a retirement home. In some of these situations the person may

be rewarded for friendliness, leadership, and energetic behavior. In others, independence, compliance, or quiet may be preferred and reinforced. After months or years in such circumstances, the individual may come to internalize the system of rewards and punishments, to assume the qualities promoted by the environment. Personality may change.

Note that this kind of change is not necessarily age related. Chance may play a large role (Bandura, 1982). What are neighbors like? What jobs are available? What does the family expect, and what are the effects of remarriage or of having or losing children? Conceivably one might be an introvert at 20, 40, and 60 and an extravert at 30, 50, and 70. Later personality might not be predictable at all from earlier personality. If there were any continuity in personality, it could be attributed to the stability of supporting environments. If the life structure remains the same, personality will too; but a change in the first could lead to a change in the second.

This extreme environmentalist position, compatible with older versions of social learning theory, is appealing neither to personality theorists nor to developmentalists. It locates the origin of behavior in circumstances, temporarily internalized, but easily replaced when circumstances change. Personality psychologists like to believe that personality is an intrinsic part of the person, changeable perhaps, but not quite so readily and with so little regard for the qualities that the individual brings to his or her exchanges with the environment. Developmentalists would also take exception to the social learning view, since they see personality as something that unfolds more or less naturally, each phase an outgrowth of earlier developments. Personality is something with history, they would say, not simply a mirror of current events.

Of course, the fact that the environmentalist view of personality is distasteful to those interested in personality and aging does not in the least mean that it is wrong. Twenty years ago, in fact, a majority of psychologists would probably have said that it was correct (many still would; see Lewis, 2001). At that time there were few studies that could really address the issue, but a number have since appeared.

True or false, the environmentalist view is certainly less beguiling than developmental theories, with their sometimes elaborate chains of growth, unfolding, and transformation. Central to the idea of development is notion of *continuity*: In any developmental sequence, a single organism goes through a series of changes, each an expression of the

same underlying entity. The butterfly is the natural outgrowth of the caterpillar, as different as the two are in form and function. The environment may hasten or retard, facilitate or damage the transformation, but caterpillars will never turn into spiders no matter what environments we impose on them. The same basic genetic material endures through and is manifest in successive stages of development.

Most of the theorizing about the course of personality in the individual has been offered by child developmentalists (J. H. Block & J. Block, 1980; Kagan & Moss, 1962; Thomas, Chess, & Birch, 1968) interested in accounting for personality from the period of infancy through adolescence, although a few theorists have carried the idea through adulthood. The attempt to account for the origins of personality in childhood is immensely attractive, since childhood seems to be the time when the most could be done to change or improve lifelong patterns of adjustment. At the same time, it is an extraordinarily ambitious undertaking. It is easy enough to chart the course of Extraversion, say, in an adult: Simply have him or her fill out a questionnaire every 10 years or so. But we cannot ask a 5-year-old to complete a questionnaire; we cannot ask an infant even a few simple questions. How, then, do we measure personality?—By observations of behavior? What kind of infant behavior corresponds to Openness to Aesthetic Experience or Dutifulness? Is it even meaningful to ask about such dimensions of personality before adolescence?

Fortunately, we do not need to solve these problems here. Other writers, more knowledgeable in the area, have struggled with them and offered their own answers. The issue concerns us here, however, because the same answers may be useful in conceptualizing the course of personality in adulthood.

One of the more elaborate models has been labeled the *heterotypic* (or *different-form*) *continuity model* (Kagan, 1971). Just as an insect goes through four distinct but developmentally related stages, so too may personality evolve through a succession of distinct phases. Sociability in adulthood may be the consequence not of sociability in childhood but of some very different trait—say, academic interests. If we watched an unsociable but intellectual child develop into a warm and friendly adult, we might believe that we had seen a failure of continuity. But if we watched a whole group of intellectual children become sociable adults, we would believe we had discovered a pattern, a kind of dispositional metamorphosis. (We might then attempt to explain the

transformation: Perhaps the social advantages of a good education lead to a more congenial environment in adulthood, making the person more friendly. Or perhaps both academic pursuits in childhood and sociability in adulthood reflect an underlying tendency to please significant others in one's life. Speculations on what might account for such relationships are endless and endlessly intriguing, which may account for the popularity of this model with many developmentalists.) Almost any trait in childhood might develop into almost any other; and of course the changes need not be limited to childhood. Young adults may develop into older adults with quite different, but quite predictable, sets of traits (Livson, 1973).

For a brief time—before additional data disillusioned us—we thought we had found just such a pattern (Costa & McCrae, 1976). In looking at cross-sectional data on several aspects of Openness to Experience, we seemed to see a shift in the aspects of experience to which open men were open. In young men, when youthful romanticism was at its peak, Openness was expressed as a sensitivity to feelings and to aesthetic experiences. In middle age, when the responsibilities of raising a family and furthering a career were uppermost, the same Openness was seen in more intellectual form, of which curiosity and a willingness to reformulate traditional values were the hallmarks. Finally, in the wisdom of age, the open person was sensitized to both feelings and thoughts, both beauty and truth. Since differentiation and subsequent integration are familiar developmental processes (Werner, 1948), this sequence seemed to provide a classic case of personality development in adulthood.

We had to abandon this model soon after we proposed it, because we found that later and better data offered no support for it whatsoever. Individuals who are open to feelings and aesthetic experiences tend also to be open to ideas and values (and actions and fantasy) at *all* ages in adulthood. This is as true for women as for men, our later studies showed (McCrae & Costa, 1983a). Age, it now appears, has nothing to do with the structure of Openness.

The heterotypic continuity model has most frequently been discussed by researchers interested in personality development in early childhood, when observable behavior (such as the smiles or cries of an infant) may bear only the vaguest resemblance to later personality traits such as Sociability or Anxiety. The model's use there may or may not be justified—a good deal more research is needed before we will know for certain.

One version, however, is familiar to students of adult personality development. As we noted in the Chapter 1, Jung (1923/1971, 1933) proposed one of the first models of adult personality development as part of his vast and intricate psychology. Instead of traits, he described various functions or structures in the psyche that governed the flow of behavior and experience. The *anima* and *animus*, for example, are parts of the self corresponding to the feminine part of the man and the masculine part of the woman, respectively. Thought and feeling, sensation and intuition are functions of the mind for perceiving and evaluating reality. The persona is the part of our personality we show to others; the shadow, the part we conceal. All these and more form the self; and the goal of adult life is the full development and expression—the individuation—of the self.

For Jung, most of these functions and structures are opposites and cannot operate simultaneously. We cannot at once be masculine and feminine, intuitive and logical, courteous and contemptuous. If all these are to be expressed, they will have to take turns. Jung proposed that this balancing process would take a lifetime; he hypothesized that the functions that dominated youth would be replaced by their opposites in old age. The aggressive, forceful young man would become docile and passive with age; the passive young woman would become aggressive. Gutmann (1970), you may recall, proposed this particular transformation as a general rule.

It is somewhat hazardous to suppose that the functions hypothesized by Jung correspond to anything as straightforward as traits, but the notion of balancing might easily be applied to them. Instead of asserting, as the heterotypic continuity model does, that any trait may give rise to any other trait, we might hypothesize that each trait will lead into its own opposite. We can then predict what an individual's personality will be like in old age by reversing the characteristics seen in youth. Introvert will be extravert, open will be closed, agreeable will be antagonistic.

Finally, there is one more model, one more possible developmental sequence: There may be no change at all. Depressed youth may become depressed elderly; talkative oldsters may once have been talkative youngsters. It may seem strange to speak of development when there is no change, so perhaps this should be distinguished as a personological rather than a developmental model. But it shares with developmental schemes the central notion of continuity. Personality, according to this theory, is not at the mercy of the immediate environment; indeed, it

withstands all the shocks of life and of aging. Catastrophic events—illnesses, wars, great losses—may alter personality, as may effective therapeutic intervention. But according to stability theory, the natural course of personality in adulthood is unchanging.

INVESTIGATING THE COURSE OF PERSONALITY TRAITS

Cross-sectional studies of personality traits tell us nothing at all about their natural histories. For that, we must trace the levels of traits through the lives of aging individuals. If we are interested in formulating general principles and not simply writing individual biographies, we must have data from reasonably large samples of individuals. And if we have enough information to draw valid conclusions, we probably have too much to be able to understand it by simply inspecting it with the unaided mind. We have to resort to statistical summaries of the data, and of these the most important for the present purpose is the correlation coefficient.

The correlation coefficient, which expresses the direction and strength of the association between two variables as a number between -1.0 and $+1.0$, is fundamental to trait psychology, being used to quantify reliability and validity and forming the basic inputs for factor analysis. It also provides the basic metric of personality stability: A stability coefficient is simply the correlation of a measure administered at one time with the same measure administered at a later time. (Note that this is also the operational definition of a retest reliability coefficient, although this latter term is generally used when the retest interval is a matter of days or weeks rather than years. We will return later to conceptual differences between the two coefficients and some implications for estimating true stability.)

Although they are familiar, correlation coefficients can be difficult to interpret. What, for example, is a *high* or *good* or *impressive* or *strong* correlation? What is moderate? What is modest or weak? These disarmingly simple questions are the source of profound and sometimes vitriolic controversy in psychology. Being human, researchers sometimes tend to consider the correlations that support their point of view “strong,” and the correlations that support other positions “weak.” Even the most unbiased judgment, however, must somehow take into account a range of considerations, including the expected magnitude of

association, the size of correlations typically found in that field of research, and the reliability of the measuring instruments.

In a few cases standards have become generally accepted. Reliability coefficients for personality tests, for example, are supposed to be in the vicinity of .70 to .90; most researchers would look askance at a test whose internal consistency was .40. In the prediction of behavior from a personality test, on the other hand, .40 would be quite respectable. One statistician, Jacob Cohen (1969), offered his expectations for psychological research in a set of rules of thumb: .10 is small; .30 is moderate; .50 or better is a large correlation. It may be useful to the reader to consider a few familiar examples. The correlation of aptitude tests with college grades is about .40–.60 (Edwards, 1954); height and weight show a correlation of roughly .70 (Peatman, 1947); and self-reports of anxiety and depression correlate about .60 in the Baltimore Longitudinal Study of Aging (BLSA).

If we obtain a set of measurements of personality traits and several years later measure the same set of people once more on the same set of traits, we would have the minimum information necessary to begin to evaluate the alternative theoretical positions we have described above. From the correlation of each trait at the first time with each trait at the second, we could test quite specific hypotheses about personality development.

If the environmentalist position is correct in positing that personality is freely reshaped by changing circumstances, there should be little or no correlation between any of the traits at the first time and any of the traits at the second. If sufficient time has elapsed to allow many individuals in the sample to move, change jobs, marry, or have children, we should expect substantial change in personality for many people in unpredictable directions. Thus, personality at Time 2 should be unrelated to personality at Time 1. We might expect modest correlations (say, .30) between corresponding traits at the two times because some individuals would have remained in the same environments, which might have sustained the traits.

The heterotypic continuity model would predict high correlations somewhere, but it might not be possible to predict just where they would occur. Extraversion at Time 1 might be strongly correlated at Time 2 with Openness or with Agreeableness or with Conscientiousness. The distinctive feature of this model is that we would *not* expect Extraversion at Time 1 to be correlated chiefly with Extraversion at

Time 2. It is principally for this reason that we would not want to conduct a study in which *only* Extraversion was measured at two times: If the heterotypic continuity model is correct, we would miss important evidence of personality continuity and transformation because we would not have assessed the trait into which Extraversion metamorphosed. By measuring all five personality factors, we are in a position to give a definitive test of the heterotypic continuity model in adulthood.

The balancing model of personality development is somewhat easier to evaluate, because we know the direction of the change we are supposed to expect. If we measure traits at the beginning of adulthood and then again at the end, we would expect a strong correlation between corresponding measurements of each trait—but we would hypothesize that the correlation would be *negative*. The extreme introvert would have become the extreme extravert; the feminine person would have become masculine; the open individual would have become closed. The predictions of this model, however, are somewhat harder to anticipate if the time span between measurements is a matter of years instead of decades, since individuals may not yet have reached the hypothetical crossover point. We would, however, expect great variability among individuals at the beginning and end of adulthood, but similarity among the middle aged, all of whom are nearing the central crossover point.

Finally, the predictions of the stability model are straightforward: If personality traits are relatively unchanging over the years, there should be high positive correlations between corresponding traits over the two times; and the correlations of each trait with itself over time should be higher than the correlations across time between different traits. In other words, the same pattern that we call retest reliability when the test is readministered after a month would be seen as stability of personality if it were readministered after a decade.

Longitudinal Evidence

Many of the studies we described in Chapter 4 as sources of longitudinal evidence about changes in the average levels of traits have also provided evidence about the course of traits in individuals. Let us begin with an examination of data from the BLSA for the 114 men who took the GZTS on three occasions about 6 years apart (Costa, McCrae, & Arenberg, 1980). Table 11 gives the observed intercorrelations among the 10 traits across 12 years from the first to the third administration.

Several things are immediately apparent from an examination of Table 11. Most obvious are the correlations, shown in boldface, that indicate the stability of individual traits. These correlations, ranging from .68 to .85, are—by Jacob Cohen's (1969) standards—extraordinarily high. Most psychologists might expect to see correlations this high if the measures were administered 12 days apart but not 12 years apart; and, in fact, the stability coefficients presented here are somewhat higher than the short-term reliabilities reported in the test manual (J. S. Guilford et al., 1976). Since the mean levels change very little (as we saw in Chapter 4), it becomes clear that most individuals obtain almost exactly the same scores on these tests on two different occasions separated by 12 years.

What about the alternative models of personality development? We need only consider the continuity models, because the data seem sharply inconsistent with any positions that do not recognize the continuity of personality. It is hardly believable that external environments have remained unchanged over the 12-year period—that would be far more puzzling than the stability of personality. Within the continuity models, the choice seems equally clear. The Jungian notion of balancing would predict *negative* correlations for the retest of traits over long in-

TABLE 11. Correlations Among GZTS Scales over 12 Years

Third administration	First administration									
	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. General Activity	.80	-.13	.38	.27	.22	.21	-.15	-.01	-.03	-.09
2. Restraint	-.13	.71	-.21	-.33	-.03	-.05	.13	.18	.01	.04
3. Ascendancy	.38	-.19	.85	.58	.33	.29	-.13	.16	.10	.04
4. Sociability	.29	-.33	.54	.75	.21	.28	-.01	-.06	.04	-.13
5. Emotional Stability	.15	-.02	.39	.35	.71	.66	.35	-.20	.37	.13
6. Objectivity	.06	-.02	.30	.30	.51	.74	.48	-.16	.37	.24
7. Friendliness	-.24	.15	-.23	-.04	.27	.41	.77	-.21	.35	.18
8. Thoughtfulness	.11	.16	.11	-.03	-.16	.18	-.28	.71	-.29	-.09
9. Personal Relations	-.12	-.05	.07	.05	.13	.25	.23	-.13	.68	.34
10. Masculinity	-.03	-.07	.11	-.04	.19	.32	.17	-.12	.24	.73

Note. Correlations greater than .24 in absolute magnitude are significant at $p < .01$, $N = 114$. Stability coefficients are given in boldface.

tervals and zero correlations for moderate intervals; the correlations in Table 11 are strongly *positive*.

The heterotypic continuity model also fails to account for the results in Table 11. True, there are some sizable correlations between traits at the first administration and different traits at the third. For example, ascendance predicts later sociability with an impressive correlation of .54. But before concluding that ascendance *develops* into sociability, we should note that the reverse is equally true: sociability predicts ascendance 12 years later with a correlation of .58. Finally (although it is not shown in Table 11), the correlation of Ascendance with Sociability when both are measured at the same time is .64 at the first administration and .58 at the third administration. In short, regardless of when either is measured, the correlation of these two traits is about .6, a fact that is not in the least surprising when one recalls that both are facets of the domain of Extraversion. In the same way, strong predictive correlations are found between different traits in the domain of Neuroticism such as Emotional Stability and Objectivity. But none of the heterotypic correlations (between one trait and a different trait) is as strong as the correlation of each trait with itself over time, and none of the predictive correlations cross the boundaries of their own domains. Extraversion does not predict Neuroticism; Masculinity does not predict Restraint.

If the data in Table 11 were the only evidence of this exceptional degree of stability in personality, we would quite properly be skeptical, but a number of other studies have reported similar findings. In the Boston study we found 10-year stability coefficients of .69 for Neuroticism and .84 for Extraversion (Costa & McCrae, 1977). More recently, Costa, Herbst, et al. (2000) reported retest correlations for more than 2,000 men and women who were initially measured in their early 40s and then retested 6–9 years later; stability coefficients ranged from .76 to .84 for N, E, O, A, and C. As long ago as 1955, Strong showed that occupational interests (which are closely related to personality dispositions) were highly stable after the age of 25.

A number of longitudinal studies by other investigators have been reported; some of them are summarized in Table 12. The correlations reported here uniformly support the stability position, although many of them are lower than those in Table 8 (in Chapter 4), for several reasons. Some subjects were college age when testing began and had not yet reached psychological maturity. Finn (1986), for example, found substantially higher retest correlations for subjects initially aged 43–53

TABLE 12. Stability Coefficients from Some Longitudinal Studies Using Self-Reports

Study	Instrument	N	Sex	Initial age	Retest interval	Correlations	
						Range	Median
J. Block (1977)	CPI	219	M, F	31–38	10		.71
Siegler et al. (1979)	16PF	331	M, F	45–70	2		.50
Leon et al. (1979)	MMPI	71	M	45–54	13	.07–.82	.50
	MMPI	71	M	58–67	17	.03–.76	.52
	MMPI	71	M	45–54	30	.28–.74	.40
Mortimer et al. (1982)	Self-concept	368	M	College	10	.51–.63	.55
Conley (1985)	KLS factors	378	M, F	18–35	20	.34–.57	.46
A. Howard & D. Bray ^a	EPPS	266	M	Adult	20	.31–.54	.42
	GAMIN	264	M	Adult	20	.45–.61	.57
Stevens & Truss (1985)	EPPS	85	M, F	College	12	–.05–.58	.34
	EPPS	92	M, F	College	20	–.01–.79	.44
Finn (1986)	MMPI factors	96	M	17–25	30	–.14–.58	.35
	MMPI factors	78	M	43–53	30	.10–.88	.56
Helson & Moane (1987)	CPI	81	F	21	22	.21–.58	.37
	CPI	81	F	27	16	.40–.70	.51
	ACL	78	F	27	16	.49–.72	.61
Helson & Wink (1992)	CPI	101	F	43	9		.73
	ACL	96	F	43	9		.73

Note. CPI = California Psychological Inventory; 16PF = Sixteen Personality Factor Questionnaire; MMPI = Minnesota Multiphasic Personality Inventory; KLS = Kelly Longitudinal Study; EPPS = Edwards Personal Preference Schedule; GAMIN = Guilford–Martin Inventory of Factors; ACL = Adjective Check List. Adapted from Costa and McCrae (1997).

^aPersonal communication, May 10, 1985.

than for those initially aged 17–25. Differences in instrument are also important. The MMPI scales studied by Gloria Leon and her colleagues (Leon, Gillum, Gillum, & Gouze, 1979) were intended to assess psychopathology, not personality traits, and so are not fully appropriate for a study of personality in a psychiatrically normal sample. Jack Block's (1977) results, using a standard personality inventory in a large group of men initially over age 30, are precisely in line with the findings in Table 11.

The 6-year longitudinal study of the NEO-PI discussed in Chapter 4 (Costa & McCrae, 1988b) was also the basis for an analysis of the sta-

bility of individual differences. Table 13 presents stability coefficients for men and women subdivided into two age groups, as well as for the total for all NEO-PI scales. (Note that the stability coefficients for A and C are based on a 3-year retest interval, because A and C were not measured until 1983.) These correlations, ranging from .55 to .87, are all significant and all testify to extraordinary stability of personality, in men and women, in early and late adulthood.

The results in Tables 11–13 provide consistent evidence of high levels of stability over intervals of up to 30 years. Consider what that means. In the course of 30 years, most adults will have undergone radical changes in their life structures. They may have married, divorced, remarried. They have probably moved their residence several times. Many people will have experienced job changes, layoffs, promotions, and retirement. Close friends and confidants will have died or moved away or become alienated. Children will have been born, grown up, married, and begun a family of their own. Individuals will have aged biologically, with changes in appearance, health, vigor, memory, and sensory abilities. Wars, depressions, and social movements will have come and gone. Most people will have read dozens of books, seen hundreds of movies, watched thousands of hours of television. *And yet, most will not have changed appreciably in their standing on any of the five dimensions of personality.*

The Time Course of Stability

Although the generalization that individual differences are preserved across time is supported by a rich body of evidence, we cannot claim that there are no changes. Quite aside from unusual cases, such as sustained traumatic experiences and effective psychotherapy, there is a certain degree of instability in the course of normal aging. The degree of stability seen is a function both of the retest interval and the initial age from which it is reckoned.

When people complete the same personality measure twice over an interval of a few days, it is not usually considered a measure of stability at all, but rather a way to assess the retest reliability of the measure, because it can be assumed that personality has not changed in so short a time. Good measures have high retest reliability, although it is never perfect, and retest unreliability sets an upper limit to the stability that can be observed.

TABLE 13. Stability of NEO-PI Scales for Younger and Older Men and Women

NEO-PI scale	Age 25–56 years		Age 57–84 years		Total
	Men	Women	Men	Women	
Neuroticism	.78	.85	.82	.81	.83
Anxiety	.74	.72	.72	.75	.75
Hostility	.74	.74	.75	.72	.74
Depression	.62	.77	.71	.66	.70
Self-Consciousness	.74	.81	.79	.78	.79
Impulsiveness	.66	.65	.70	.67	.70
Vulnerability	.73	.76	.60	.75	.73
Extraversion	.84	.75	.86	.73	.82
Warmth	.74	.66	.74	.66	.72
Gregariousness	.76	.69	.76	.68	.73
Assertiveness	.73	.75	.84	.79	.79
Activity	.72	.74	.77	.74	.75
Excitement Seeking	.69	.61	.79	.63	.73
Positive Emotions	.80	.65	.69	.69	.73
Openness	.87	.84	.81	.73	.83
Fantasy	.77	.74	.69	.63	.73
Aesthetics	.75	.83	.81	.67	.79
Feelings	.72	.61	.64	.62	.68
Actions	.75	.73	.66	.58	.70
Ideas	.78	.84	.78	.75	.79
Values	.77	.69	.70	.67	.71
Agreeableness	.64	.60	.59	.55	.63
Conscientiousness	.82	.84	.76	.71	.79

Note. All correlations are significant at $p < .001$. N 's = 63 to 127 for subsamples. Retest interval is 6 years for N, E, and O; 3 years for short A and C scales. Adapted from Costa and McCrae (1988b).

Other things being equal, the longer the retest interval, the lower the stability coefficients. In Table 12, a study by Leon et al. (1979) is reported. Here the same individuals were retested twice, once after 13 years, and again after another 17 years. The median stability coefficient after 13 years was .50; across the whole 30-year interval it was only .40. Our own studies of the GZTS continued after the 12-year study reported in Table 9; by 1992 we had GZTS data over a 24-year period. The median 12-year stability was .74, whereas at 24 years the median stability coefficient had declined to .65—still amazingly high, but clearly lower. This generalization is supported by a meta-analysis (a quantita-

tive review of the literature) conducted by Roberts and DelVecchio (2000). Small changes accumulate here and there over time, and the net results is a slow decay of initial rank ordering. Based on GZTS data, we projected that about 60% of the variance in personality trait scores is preserved over the full adult lifespan (Costa & McCrae, 1992b).

The degree of stability is also related to the initial age of the sample, as Roberts and DelVecchio (2000) showed. Costa, Parker, and McCrae (2000) reported stability coefficient for 12-year-olds retested at age 16. These ranged from .30 to .63 with a median of .38. R. W. Robins, Fraley, Roberts, and Trzesniewski (2001) reported on changes over the 4 years of college; these somewhat older adolescents were already more stable, showing stability coefficients ranging from .53 to .70 (median = .60). Contrast both of these with the values that Costa, Herbst, et al. (2000) calculated for 40-year-olds retested after 6–9 years, where the median across the five personality factors was .83. Adults show much higher stability than early adolescents and substantially higher stability than college students.

At what age does personality “stabilize,” that is, reach the maximum degree of stability for a given retest interval? That is a hotly contested question. Based on their meta-analysis, which statistically combined every study they could locate in the literature—some 3,000 correlations in all—Roberts and DelVecchio estimated that stability coefficients do not peak until age 50–60 and predicted that retest correlations from 40-year-olds should be about .60 over a 7-year interval. They argued from this that there is more flexibility in adulthood and more possibility for growth and change than our earlier writings had suggested.

But there are reasons to be skeptical of these meta-analytic results. After all, we know from studies of 40-year-olds that the observed stability is much higher than .60, at least when the NEO-PI-R is the instrument used (Costa, Herbst, et al., 2000). It turns out that most of the studies Roberts and DelVecchio had to review were of children and adolescents, which obviously do not bear directly on the issue of stability in adulthood. And meta-analyses are by design nonevaluative—they combine good and poor studies equally—and by necessity are limited to the studies that have been published. If there are only a few studies in the literature and some of them are suboptimal, then no amount of statistical manipulation can arrive at trustworthy results.

In meta-analyses results are pooled from studies with different

samples and different measures, conducted at different times. The hope is that all these influences will balance out, leaving a good estimate of the true value. But there is another, arguably better, way to resolve the question of when personality stabilizes: Compare retest correlations for younger and older samples measured at the same times and with the same instrument. We have done that with the GZTS (Costa et al., 1980) and with the NEO-PI (Costa & McCrae, 1988b), and we have consistently found that men and women in their 30s and 40s show stability coefficients as high as men and women in their 50s and 60s. Table 13 illustrates this fact: The median stability coefficients for the younger groups are .74 for both men and women; for the old group, they are .75 for men and .69 for women. Already at age 30, men and women appear to show extremely high consistency over time in their personality traits.

* * *

The idea that personality should be so deeply and permanently ingrained was revolutionary in a scientific climate in which change, growth, and development were the watchwords. Not surprisingly, therefore, a number of psychologists challenged the basic findings when they were first published.

When unanticipated results are found, there is a strong tendency to dismiss them. This reluctance to accept an unforeseen result is not really a result of closed-mindedness, nor is it unscientific. On the contrary, the essence of science is to look critically at observations and conclusions. Every scientist knows that there is more than one explanation for any phenomenon and is rightfully wary of accepting the first account that offers itself. True, the high correlations between personality scores at two different times are just what one would expect if personality is actually highly stable. But there are a number of other ways in which the same correlations could be observed that are entirely consistent with major change in personality. Until these rival hypotheses can be ruled out, it is premature to accept the conclusion of stability.

Stability Reconsidered

Qualifications and Rival Hypotheses

METHODOLOGICAL ISSUES IN THE ASSESSMENT OF STABILITY

The first and most fundamental challenge to any interpretation of data is the notion of chance: Perhaps the high stability coefficients were flukes. In this particular case that argument is extremely weak. Statistically, the likelihood of observing correlations of this size in samples this large purely by chance is less than 1 in 10,000. And it would be coincidence indeed if we happened to exceed these odds for measures of all five domains of personality. No scientist would seriously suggest these are chance results.

If the only data were from the study of the Guilford–Zimmerman Temperament Survey (GZTS) in the BLSA, one might argue that the results are not generalizable to other segments of the population. Women were not included in that study; Blacks and other non-Whites were greatly underrepresented; education, social class, and intelligence were markedly higher in the BLSA study than they are in the full population: Perhaps these characteristics somehow explain the observed stability. However, our conclusions are not based on Table 11 (in Chapter 6) alone. Similar evidence comes from a variety of studies of men and women. There is strong replicability of results over samples that are considerably more diverse and representative than the BLSA sample.

The same argument—and the same answer—applies to consider-

ations of measures. If stability were found only when the GZTS was used, we would begin to suspect that GZTS scores, not personality dispositions, were stable, and we would then try to determine what peculiarity about the GZTS accounted for its constancy. But highly similar results have been seen with the I6PF and the NEO-PI; in fact, almost every personality inventory that has been longitudinally examined has shown exceptional levels of stability.

And so the attack shifts ground. If it is not a specific personality test that accounts for stability, perhaps it is the nature of personality tests themselves. All of the instruments examined in Chapter 6 in Tables 11–13 are self-report inventories, in which we base our assessment of the individual on the ways in which he or she responds to a series of more or less straightforward questions. It has been known for some time that the scores we obtain in this way are not pure indicators of personality. Instead, any of several *response sets* may be operating, and these are likely to cloud the issue (Jackson & Messick, 1961). Some individuals agree to almost any characterization of them you offer; some reject them all. The former are called *yea-sayers*, or acquiescers; the latter, *nay-sayers*. Some people are extreme responders: They tend to express their beliefs about themselves and other things strongly. Other people are more noncommittal and frequently use the “neutral” or “don’t know” response category. Social desirability has been the bane of many test constructors: How do we know that individuals are not lying in order to make themselves look good? How do we know that respondents are not fooling themselves as well as us?

Other response sets have also been noted: tendencies to respond carelessly, to try to make a bad impression (out of a kind of test-taking perverseness), to answer consistently. Study after study has documented the potential threat that each of these may pose to a straightforward interpretation of data, and scales have been incorporated into many prominent personality inventories in order to screen out the worst offenders. From time to time critics have alleged that paper-and-pencil tests may be explainable completely in terms of these stylistic (and substantively insignificant) sets (Berg, 1959), but careful analyses have consistently shown that this “nothing but” interpretation is far too extreme. Still, it is possible that the very high levels of stability in Tables 11 and 13 are inflated by response sets.

Consider first the possibility that subjects are motivated to respond consistently from administration to administration: They recall how

they answered questions the first time and repeat their performance on subsequent administrations despite real changes in their personality. This possibility is frequently advanced, but it becomes more and more implausible on examination. Why would people be so concerned with maintaining the appearance of consistency? A few subjects might be obsessed with maintaining an image, but it is hard to believe that nearly everyone in the sample would be (though that would be necessary account for the observed coefficients). Further, even if they wanted to, is it likely that subjects *could* repeat their earlier performance? Informally, subjects often report that they do not even remember taking the test before—are we to believe that they somehow recall how they answered each particular question?

This deliberate consistency hypothesis is almost too implausible to entertain, so it is fortunate for us that it has already been tested. Woodruff and Birren (1972) retested a sample of individuals who had completed a personality measure 25 years before when they were in college. In addition to being given a simple readministration, the subjects were also asked to fill out the questionnaire *as they believed they had filled it out 25 years earlier*. In terms of average level, the original responses were closer to the readministration responses than they were to the subjects' recollections of their original responses. Woodruff (1983) demonstrated this even more forcefully when she examined the correlations between real and recalled responses. The correlation between original scores and later recollections was low and statistically nonsignificant. By contrast, there was a strong and highly significant correlation between original responses and personality as measured 25 years later. We get a more accurate picture of what people were like 25 years ago by asking what they are like today than by asking them to recall how they were; memories, it seems, are not as stable as personality.

A more plausible interpretation of a response set explanation would be based on the following analogy: When individuals take a test, they employ a characteristic set of response styles that operate much like habits. Just as one never forgets how to ride a bicycle, so one may never forget habits of test responding. When subjects retake the test years later, the same habits are activated and the same results obtained. Thus, there would be stability, but it is stability of response style, not personality.

This is not as strong an argument as it may at first seem. It might account for the convergence between scores on the same trait at two

times, but it cannot explain the divergence between scores on different traits. The scales of the GZTS differ much more in content than in the response styles they are likely to elicit. Thus, the low correlations between unlike traits (e.g., sociability and thoughtfulness, $r = -.06$) are easily explained if content determines responses but are hard to explain if response sets determine responses. And if response sets do not account for scale scores at one time, they cannot account for the stability of scale scores across time. It is, however, still possible that stability is exaggerated by consistency of response habits.

Because response sets (including random responding, uncertainty, social desirability, and acquiescence) can be measured in the GZTS, we conducted empirical tests of this possibility (Costa, McCrae, & Arenberg, 1983). We found that there is indeed a certain consistency of response style across administrations, but it is less pronounced than the consistency of individual traits. And when we analyzed the data to eliminate the effects of this consistency, we found that the stability coefficients remained essentially unchanged. Artifacts of response set do not account for stability of personality.

The Self-Concept and Stability

A much more interesting and reasonable rival hypothesis was suggested to us by Seymour Epstein. He drew attention to the distinction between personality and the self-concept, between what we are really like and what we believe we are like. He suggested that we may develop a view of ourselves early in adulthood and cling to that picture over the years despite changes in our real nature. Rosenberg (1979) expressed the same idea this way:

The power and persistence of the self-consistency motive may be quite remarkable. People who have developed self-pictures early in life frequently continue to hold to these self-views long after the actual self has changed radically. . . . [T]he person who grows gruff and irritable with the passing years may still thank of himself as "basically" kindly, cheerful, and well-disposed; thus behavior which has become chronic is either unrecognized or is perceived as a temporary aberration from the true self. (p. 58)

Psychologists and sociologists have long known that we do have a self-concept, a theory of what we are like (Epstein, 1973), and that it seems to guide our behavior. It is this theory of ourselves that we draw

on when we are asked to describe ourselves, whether to a friend or on a questionnaire. There has been considerable dispute about how we develop a self-concept, whether or not self-concepts are accurate, and what the possibilities are for change in the self-concept during adulthood. According to one view, the self-concept may be crystallized early in life and remain unchanged thereafter.

Stories abound about middle-aged men who take up tennis or jogging after 20 years of inactivity and have to revise drastically their image of themselves as athletes. Normally they make these revisions fairly quickly. They have to; they can't fool their bodies. But could they still imagine they were adventurous when they had lost their sense of daring? Could they still believe they were hostile when they had been mellowed by experience and age? Would anything in their intrapsychic or social experience *force* them to revise their self-concept, the way their exhaustion forces them to acknowledge the limits of their physical endurance?

Until we can answer these questions, the possibility remains that all that we have demonstrated is the stability of the self-concept. Further, this critique would apply equally to all studies, longitudinal and cross-sectional, that have relied on self-reports. In short, it would pull the rug out from under most of the claims of stability we have reviewed in this book and would allow the proponents of change to say, "See? People do change with age and with experience, just as we've told you. Only they don't *realize* they've changed."

In this case we cannot rely on evidence from self-reports, since they are themselves in question. We have to get beyond self-reports, which are determined by the self-concept, if we want to see the agreement between the self-concept and the true personality. For this we must turn to someone else's assessment of personality, and in our case we had available personality ratings made by the husbands and wives of 281 of our subjects (McCrae & Costa, 1982). We assumed that spouses were reasonable judges of personality (we had data to support that assumption), but we also argued that the spouses would be much more able than the individuals themselves to detect any changes that age had brought. We may be blind to changes in ourselves, but others who must deal with us on a daily basis are unlikely to overlook important changes in our personality.

By comparing self-reports with spouse ratings, we can get some idea of whether the self-concept has really been crystallized. Among

young people the self-concept should still be quite accurate, a good reflection of what the person is really like. Consequently there should be good agreement between self-reports and spouse ratings among young couples. But if personality changes with age while the self-concept is frozen, self-reports will give increasingly inaccurate pictures of personality. Correspondingly, the agreement between self-reports and spouse ratings will become poorer and poorer. Among older couples there may be no agreement at all.

These would be the consequences if personality changed while the self-concept did not. When we examined the correlations between self and spouse for young, middle-aged, and elderly couples, however, we found no decrease in agreement with age. (In fact, the only significant difference was that there was more agreement—not less—about Extraversion in older men.) This finding provides some powerful evidence that personality does not change and that crystallization of the self-concept cannot itself account for the correlations seen in studies of self-report inventories.

There are other pieces of evidence as well. Jack Block (1971, 1981) has reported on a remarkable series of longitudinal studies conducted in Berkeley. Personality records have been kept on about 200 boys and girls over a period now approaching 50 years. These records include test scores, teachers' notes, biographical sketches, and many other documents. For each of four time periods (junior high school, senior high school, the 30s, and the 40s) independent judges reviewed these records and rated the subjects on the California Q-Set (CQS), which, as we have seen, measures aspects of all five domains of personality. Different judges were used for each time period, and yet, when individuals were compared across time, remarkable evidence of stability in personality was found.

In another study on the parents of these subjects, judges provided ratings across an interval of 40 years, from ages 30–70 (Mussen, Eichhorn, Honzik, Bieber, & Meredith, 1980). Again, despite the differences in judges and the unreliability of single-item scales, most of the ratings showed significant correlation across time.

For the domains of N, E, and O we have 6-year longitudinal spouse ratings for 167 subjects. Table 14 gives stability coefficients for men and women separately and combined; all these correlations are significant and quite comparable in magnitude to those found in studies of self-reports. Husbands' and wives' views of their spouses' personalities con-

TABLE 14. Stability Coefficients for Spouse Ratings of Men and Women

NEO-PI scale	Men	Women	Total
Neuroticism	.77	.86	.83
Anxiety	.67	.82	.75
Angry Hostility	.76	.81	.78
Depression	.69	.74	.72
Self-consciousness	.65	.77	.76
Impulsiveness	.70	.81	.75
Vulnerability	.62	.70	.68
Extraversion	.78	.77	.77
Warmth	.76	.70	.75
Gregariousness	.71	.72	.73
Assertiveness	.68	.75	.72
Activity	.72	.63	.68
Excitement Seeking	.65	.75	.69
Positive Emotions	.78	.77	.77
Openness	.82	.78	.80
Fantasy	.73	.73	.73
Aesthetics	.83	.70	.79
Feelings	.71	.65	.70
Actions	.78	.69	.75
Ideas	.75	.72	.75
Values	.81	.72	.76

Note. All correlations are significant at $p < .001$. Adapted from Costa and McCrae (1988b).

firm the essential stability of personality. In a later study we looked at the 7-year stability of peer ratings of men and women (Costa & McCrae, 1992b). For the five NEO-PI domains, the retest correlations ranged from .63 to .84.

Critics could, of course, argue that spouses and peers form a crystallized concept that is the source of apparent stability, but such arguments become increasingly implausible. It is far more parsimonious and plausible to admit that personality really is stable.

Accounting for Variance and Correcting for Unreliability

The typical 10-year stability coefficient of .71 reported by Block (1977) can be interpreted to mean that 50% (or .71 squared) of the variance in

scale scores at Time 2 can be predicted by Time 1 scores. Some critics of the stability position have asked about the remaining 50%: Isn't it fair to say that there are equal amounts of stability and change?

This argument would be sound if personality scales were perfect indicators of personality. But even the best instruments are fallible. When a subject encounters the NEO-PI item, "I really like most people I meet," he or she must choose among five options, from "strongly disagree" to "strongly agree." If the subject happens to recall meeting several obnoxious people lately, he or she may respond "disagree." On a second administration, the subject may place more weight on his or her generally friendly response to people and "agree" with the item. The subject's personality has remained the same, but his or her reading of the item has changed. It is for this reason that scales combining many items are generally more accurate than those that include only a few: Random errors tend to cancel out when many items are used. (An illustration of this principle is seen in Table 13 [in Chapter 6] and Table 14, where the longer domain scales show higher stability coefficients than their component facet scales.)

There are also other sources of error in filling out questionnaires. Respondents may be living through a difficult period, perhaps looking for work or adjusting to a divorce (Costa, Herbst, et al., 2000); their temporary mood may color their perception of themselves. Problems with physical health may also affect responses. Whatever the causes, test scores tend to fluctuate from day to day for reasons that are unrelated to personality change. Traditionally, scales are evaluated in part by their ability to give relatively constant or *reliable* scores despite these influences. This property is assessed by repeating the measure a few days or weeks later for a sample of people and correlating the first and second set of scores. Since it can be assumed that no true personality change has occurred in the interval, retest correlations less than 1.0 can be interpreted as evidence of unreliability. Personality scales normally have retest reliabilities in the range of .70 to .90.

In longitudinal studies, the correlations between the initial and later administrations of the test are less than 1.0, lowered both by the day-to-day fluctuations of questionnaire responses and by whatever true change has occurred. Retest reliability puts an upper limit on stability: We cannot expect to find high correlations over a period of years if we do not observe them over a period of days. It is therefore reasonable to ask how close to this upper limit stability coefficients reach and

what proportion of possible stability they show. Dividing stability coefficients by short-term retest reliabilities gives an estimate of the stability of the *true score*, the actual level of the traits that we try to infer from our tests.

This procedure—called *disattenuating correlations*—leads to considerably higher estimates of stability. When corrected for attenuation resulting from unreliability, 12-year stability estimates for the GZTS scales shown in Table 11 (in Chapter 6) ranged from .80 to 1.00 (Costa et al., 1980); the 6-year stability estimates for N, E, and O scales for the NEO-PI were .95, .90, and .97, respectively (Costa & McCrae, 1988b). Such analyses suggest that if we had perfect measures we would find near-perfect stability of personality traits. The 50% of variance not predictable from earlier scores is not true change; it is mostly error of measurement.

RETROSPECTION AND SELF-PERCEIVED CHANGE

In Chapter 6 we mentioned that retrospection—asking individuals to recall how they were and how they have changed—was a simple but suspect method of examining stability or change in personality. We said that we could not be sure whether the reports would be accurate reflections of reality or distortions of memory. By now, however, we have a reasonable idea of what reality must have been like for most people: We know that stability rather than change predominates when prospective longitudinal studies are conducted. So is there any reason to reconsider retrospective reports?

In fact, there is. They serve, to begin with, as one more source of data, one more way to bolster or cast doubt on our findings of constancy. But equally importantly, they let us know something about how individuals view their own lives. If most individuals were to claim that they had changed dramatically when all the evidence points to the opposite conclusion, this would suggest that memory does indeed play tricks on us, and the function of these tricks would be of considerable interest. It might also explain why critics find the notion of stability in personality so counterintuitive.

Most people, in fact, can point to aspects of themselves that have changed in the past few years, and the reader may have been protesting on the basis of these exceptions throughout this book. But three things should be borne in mind in considering this kind of evidence:

First, in one important sense retrospection is different from any of the kinds of studies we have reviewed. Statistics are usually interpreted in terms of individual differences; that is, they compare one individual with other. Saying that a person is warm means, in a certain sense, that he or she is warmer than most other people; saying that warmth does not change with age means that he or she will continue to be warmer than most people throughout life. But if we ask an individual if he has changed, he will be making comparisons between himself as he was and himself as he is now. He is likely to be far more sensitive to these changes than we are. We might find that he has risen from the 55th to the 58th percentile in warmth—a change we could call trivial. From his perspective, however, this change may be extremely important and the contrast may be very vivid—just as a dieter might be impressed by a loss of 5 pounds that no one else would notice.

Second, as students of visual perception would remind us, our attention is attracted by movement and contrast, not by stability and sameness. If a person is unchanged in characteristic levels of Anxiety, Hostility, Assertiveness, Excitement Seeking, and Openness to Feelings and Values but has changed in Gregariousness, he or she is much more likely to notice and remark on the change in that one element than on the stability in the other six. Our prediction of stability would be confirmed in 85% of the cases, but we would be judged wrong because of the one exception. Stability is not absolute, but it is far more pervasive than many people realize.

Finally, the question of sampling arises again. Readers under age 30 probably *have* experienced some personality changes in recent years, but they should not assume that this pattern will continue. Regardless of age, readers of this book (and lifespan developmental psychologists) are probably not representative of humanity in general. They are, for the most part, intelligent, perceptive, and curious individuals. And almost certainly they are interested in seeing what changes age and experience will bring to them. In short, they begin with a bias toward finding change and the mental acuity to spot it, no matter how subtle or small it may be. Most people are neither predisposed nor skilled enough to recognize such changes.

We first discovered this when we began to investigate the so-called midlife crisis (Costa & McCrae, 1978). At that time we assumed we would find abundant evidence of a crisis; we simply wanted to document it. (See Chapter 9 for a fuller discussion of this study.) At the end of a standard questionnaire we asked that subjects comment in their

own words on the ways in which they felt they had changed in the past 10 years. Subject after subject returned the questionnaire with words to the effect of “no changes worth mentioning.” A few subjects did find some change to comment on; the great majority did not.

Gold, Andres, and Schwartzman (1987) recently came to the same conclusion in their study of self-perceived change. They administered the Eysenck Personality Inventory to 362 elderly men and women. One month later they readministered the questionnaire to half the sample as a control group; the other half—the experimental group—were asked to discuss their life and circumstances at age 40 and then to complete the questionnaire to describe their personality at age 40. The responses of the experimental group suggested that they perceived themselves as having been slightly more extraverted at age 40 than currently, but individual differences were highly stable. Recalled Neuroticism correlated .79 with current Neuroticism in the experimental group; the 1-month reliability observed for the control group was .82. The corresponding values for Extraversion were .75 and .79, respectively. By the logic of correcting for attenuation, we could argue that these subjects perceived about 95% stability between ages 40 and 75.

Most people, then, do not see much change in their own dispositions in adulthood, but it is possible that the minority who do perceive differences are correct: Perhaps they really have mellowed or matured with age. To test this possibility we conducted a small study of self-perceived change in conjunction with our 6-year longitudinal study of self-reports and spouse ratings on the NEO-PI (Costa & McCrae, 1989b). At the end of our packet of questionnaires we asked subjects to:

Think back over the last 6 years to the way you were in 1980. Consider your basic feelings, attitudes, and ways of relating to people—your whole personality. Overall, do you think you have (a) changed a good deal in your personality? (b) changed a little in your personality? or (c) stayed pretty much the same since 1980?

We found that a bare majority (51%) believed they had “stayed pretty much the same” and another third (35%) thought they had “changed a little.” But a substantial minority, 14%, felt that they had “changed a good deal” in personality. These individuals would probably take exception to our conclusions about stability.

Are these perceptions veridical, or are they distorted by tricks of

perception or memory? One way to examine this question is by comparing stability coefficients within the three perceived change groups. If perceptions of change are accurate, we should see the highest coefficients in the “stayed pretty much the same” group and the lowest coefficients in the “changed a good deal” group. We had both self-reports and spouse ratings to test this hypothesis. But Table 15 provides no support for the hypothesis: None of the five personality factors is consistently less stable among individuals who believed they had “changed a good deal.”

In addition to the global change item we also asked subjects, “Specifically, compared to how you were 6 years ago, how lively, cheerful, and sociable are you now?” This question sought to assess each subject’s perceived change in Extraversion. Similar items asked about the other four dimensions, and subjects could respond with “more,” “less,” or “same.” We found that subjects who believe they were more extraverted or neurotic actually did score slightly higher on these two scales in 1986, but spouse ratings did not confirm the changes; perhaps there was a small but real change in self-concept, although not in true personality. There was no evidence from either self-reports or spouse ratings to substantiate self-perceived changes in O, A, or C. We later replicated this study in a much larger sample, with essentially the same results

TABLE 15. Stability Coefficients within Perceived Change Groups

NEO-PI scale	Perceived change		
	“A good deal”	“A little”	“The same”
Self-reports			
Neuroticism	.77	.79	.86
Extraversion	.66	.80	.86
Openness	.86	.83	.82
Agreeableness	.70	.56	.62
Conscientiousness	.80	.71	.81
Spouse ratings			
Neuroticism	.86	.79	.84
Extraversion	.79	.79	.72
Openness	.90	.88	.71

Note. Ns = 36–182 for self-reports, 13–71 for spouse ratings. All correlations are significant at $p < .01$. Adapted from Costa and McCrae (1989b).

(Costa, Herbst, et al., 2000). For the most part, it appears that perceptions of change in personality are misperceptions. Most people think they are stable, and those who think they have changed are probably wrong.

With the benefit of hindsight and the results of many longitudinal, prospective studies, it is interesting to consider results from an early retrospective study. Reichard, Livson, and Peterson (1962), in one of the seminal books on personality and aging, interviewed a number of retired men to find out how they were adapting to old age and retirement. The interviewers spent a good deal of time taking life histories of their subjects so that they could note patterns of adjustment across the lifespan. Their conclusion: "The histories of our aging workers suggest that their personality characteristics changed very little throughout their lives" (p. 163).

MODERATORS OF CHANGE

We have not yet pinned down the case for stability. It seems clear that most people do not change much and that those who think they have are usually mistaken. But perhaps there are subsets of people who really change but don't necessarily perceive it themselves. If people's perceptions of change can be wrong, their perceptions of stability can be, too. With theoretically based hypotheses to guide our search, we might be able to identify moderators of change—circumstances or conditions that promote growth or decline in a subgroup of people. Using the same before-and-after design that we described for the self-perceived change study, we could test these hypotheses, and a number of studies have done so.

Psychological Characteristics

One approach is to look for those characteristics of people that might make certain individuals susceptible to change. McCrae (1993) examined several of these. One obvious candidate is personality itself—especially the dimension of Openness to Experience. Some people are much more willing to explore their world, to listen to different views, and to reflect on their own feelings. Perhaps this continued processing of new information might lead to changes in emotional maturity or patterns of social interaction, or dedication to goals—to changes in personality

from which the closed individual would be insulated. The changes might be in either direction: Some open people might discover reasons to become more sociable, whereas others might find new meaning in solitude. The former would become more extraverted; the latter, more introverted. We would not, therefore, expect an overall increase or decrease. But we would expect that people high in Openness would be less predictable over time: Retest correlations should be lower for them than for people low in Openness.

The 6-year N, E, and O data and the 3-year A and C data gathered in the BLSA in the 1980s provided the basis for a test of that appealing hypothesis (McCrae, 1993). The median retest correlation for the closed half of the sample was .79, showing the high stability predicted for this group. But the median retest correlation for the open half of the sample was .80, demonstrating equally high stability of personality. This analysis offers no support for the hypothesis.

McCrae (1993) used the same design with three other variables that were potential moderators of change. *Private self-consciousness* (Fenigstein, Scheier, & Buss, 1975) refers to a tendency to be introspective and interested in understanding oneself. (This construct must be distinguished from neurotic self-consciousness, which refers to feelings of discomfort in social situations.) Because people high in private self-consciousness know themselves well, they may be more consistent in their self-descriptions and thus show higher stability coefficients. *Personal agency* (Vallacher & Wegner, 1989) is a cognitive construct that concerns the way people think about their own behavior. People high in personal agency view behaviors as steps toward long-term goals, whereas people low in personal agency take a more immediate and concrete view. For example, the act of making a will might be seen by a person high in personal agency as facing one's mortality and taking care of one's long-term responsibilities, whereas it might be seen by a person low in personal agency only as signing a paper. Because personality questionnaires ask about relatively abstract and enduring traits, they may be answered more consistently over time by people high in personal agency. *Self-monitoring* (Snyder, 1979) is the tendency to watch what one is doing and modify it to fit the situation. At best, high self-monitors are socially sensitive and functionally flexible people; at worst, they are other-directed chameleons without any enduring principles of their own. Because they can change to suit the occasion, high self-monitors might be expected to be lower in trait stability.

In brief, none of these hypotheses was supported. When divided

into high and low groups on each of these three variables, stability coefficients for the five personality factors were essentially identical. There are, of course, many other variables that have not yet been examined, but to date no psychological characteristic has been shown to moderate the stability of personality traits.

Life Events

Perhaps a more likely candidate for a moderator of change is some external event that alters one's life and thus one's personality traits. Costa, Herbst, and colleagues (2000) conducted exploratory analyses based on this premise. In their large midlife study, they asked participants to report which of 30 specific life events they had experienced in the past 6 years. Events included *got married*, *your parent died*, and *your spouse was fired*. Individual events were relatively rare in this sample, so an index of total life events was created by summing across the 30 events. Contrary to the hypothesis, there was no evidence that more events led to more change in personality traits from before the 6-year period to after it.

However, analyses of selected events did show small effects. A small group of people (4% of the sample) were fired from a job; compared to the group of people who were promoted during the same period, these individuals increased in Neuroticism and decreased in Conscientiousness. Again, women who were divorced became a bit higher in Extraversion and Openness compared to women who got married. It remains to be seen how replicable these findings are and how enduring the changes may be.

Neyer and Asendorpf (2001) recently examined changes over a 4-year period in a sample of young German adults. It is well known that there are changes in the mean levels of personality traits in this developmental era; what is not known is whether those changes are a natural and more or less inevitable effect of aging itself or whether they result from the social changes that come with adult roles of work and family. Neyer and Asendorpf hypothesized that settling down with a life partner would enhance emotional maturity and lead to lower Neuroticism, and that having a child would make individuals more responsible and thus higher in Conscientiousness.

Over the 4 years of their study, some participants remained single whereas others who were initially single became involved in a stable relationship. Consistent with hypotheses, those who remained single did

not change much in personality, whereas those who had formed a relationship had also matured in personality, showing lower levels of Neuroticism and higher levels of Conscientiousness. Note, however, that the causal ordering is ambiguous, because we do not know when in the course of 4 years the personality changes occurred. It is possible that some people spontaneously became better adjusted and in consequence found it easier to establish a lasting relationship.

Neyer and Asendorpf (2001) also looked at the transition to parenthood. From an evolutionary perspective, it would make sense to argue that having a child should trigger growth in Conscientiousness. It would also make sense from a social learning perspective, because there are many social norms that operate to encourage responsibility in parents. But although men and women did in general become more conscientious over the 4-year period, that change was unrelated to parenthood. Perhaps we have evolved under the assumption that everyone would have children at this age.

The search for personality change associated with life circumstances and events has been pursued most vigorously by Ravenna Helson, Brent Roberts, and their colleagues (e.g., Helson, Pals, & Solomon, 1997; Roberts & Helson, 1997). Many of the studies have been conducted as part of the Mills study, an intensive and long-term follow-up of about 100 women from Mills College, an elite institution in Oakland, California. Replications in larger and more diverse samples are certainly needed, but Helson and Roberts have clearly managed to keep this question open.

Physical and Mental Health

There is one last category of potential moderators of personality change: health status. The popular press has frequent accounts of people who claim to have changed their lives as a result of a brush with death or an acquired disability. Do heart attacks teach hard-driving executives to slow down and enjoy life? Can loss of hearing make a closed-minded dogmatist more receptive to others' points of view?

Appealing as this notion might be, there is already reason to be skeptical. Nothing is more characteristic of the aging process than the deterioration of physical health. Most 80-year-olds have a history of health problems and some degree of disability. Yet, as we have seen, stability coefficients for personality traits do not decline with age.

There is, however, a more direct test of this hypothesis. In collaboration with a BLSA physician, we examined the medical histories of men and women on whom we had before-and-after personality measures (Costa, Metter, & McCrae, 1994). Initially, 153 volunteers were considered healthy, 93 had a minor disease (e.g., treated hypertension), and 27 had a major disease (e.g., a history of cancer). Over time, 175 volunteers stayed the same or improved in health whereas 98 had worse health status. These two groups did not differ in longitudinal changes in any of the five personality traits, and retest correlations were equally high in both groups. Deterioration in physical health status was unrelated to change in personality in this group. It should be noted, however, that all of the subjects in this study were sufficiently well at the second administration to complete a personality questionnaire. More serious illness might have affected personality scores.

In 1991, Siegler and colleagues offered the first clear evidence of personality change associated with disease. They asked caregivers of mildly to moderately demented patients to rate the patient's personality as it had been before onset of the disease (chiefly Alzheimer's) and as it was now. They used the same version of the NEO-PI that we had administered to spouses and peers of BLSA participants. Dramatic effects were seen. Prior to their illness, the patients had not differed from the general population. But subsequently they showed increased Neuroticism, especially Vulnerability, decreased Extraversion, and a precipitous decline in Conscientiousness. Planning, organization, and self-discipline appear to be dramatically impaired by dementing disorders.

This finding has been widely replicated (Chatterjee, Strauss Smyth, & Whitehouse, 1992; Siegler, Dawson, & Welsh, 1994; Welleford, Harkins, & Taylor, 1995), and it has been extended by Trobst (1999) to a different category of neurological patients, those with long-term consequences of traumatic brain injury (chiefly from automobile accidents). Patients with Parkinson's disease show some of the same changes (Mendelsohn, Dakof, & Skaff, 1995).

Taken together, these two lines of research make a clear contrast. Diseases of the heart, or liver, or lungs may have substantial effects on people's lives, but they do not have permanent effects on personality traits. Diseases of the brain, however, often do. This deep and selective tie of personality traits to the brain is a point to keep in mind when, in a later chapter, we turn to a theoretical account of personality itself.

It also provides a new perspective on a phenomenon noted many

years ago. If patients who are clinically depressed complete a personality inventory, they tend to score extremely high on measures of Neuroticism. After successful treatment or after spontaneous remission, these patients typically score substantially lower on Neuroticism—although still well above average. One interpretation of this change in personality scores is that the acute state of depression distorts the measurement of personality, giving exaggerated, state-influenced Neuroticism scores (Hirschfeld et al., 1983). But we now know that clinical depression is a brain disease; we also know that brain diseases can alter personality traits. From this perspective it can be argued that personality assessments during depressed states are not distorted; they reflect a real, if temporary, increase in Neuroticism. Indeed, one might say that an acute increase in the level of Neuroticism is one of the major symptoms of clinical depression.

CASE STUDIES IN STABILITY

The same individuals must be studied at least twice in order to calculate the stability coefficients on which we have focused so much attention. But these coefficients are still group statistics, showing the extent to which the same ordering of individuals along a dimension is maintained over time. A less scientific but perhaps more easily grasped way to examine stability or change is by comparing the personality profiles of individuals at two time periods. Figures 5–7 provide examples of individuals in our studies tested in 1980 and again in 1986 (Costa, 1986): The 1980 scores are joined by broken lines; the 1986 scores, by solid lines.

Figure 5 gives the NEO-PI profile of a woman (Case 1) who described herself as a homemaker; she was 62 when she completed the second administration. The first five columns of the profile show her standing on the five domains; we can see that she is average in Neuroticism and Agreeableness, low in Extraversion and Openness, and high in Conscientiousness (A and C were measured only in 1986). Her facet scores are given in the next three sets of columns. It is clear that her scores are almost identical on the two occasions; even the relatively fine distinctions within facets are preserved, most strikingly for Extraversion. The only notable change is an increase from average to very high in Openness to Values. Does this represent a temporary change in attitudes or a permanent rethinking that may lead to greater openness

Females — Form S (Self-Report)

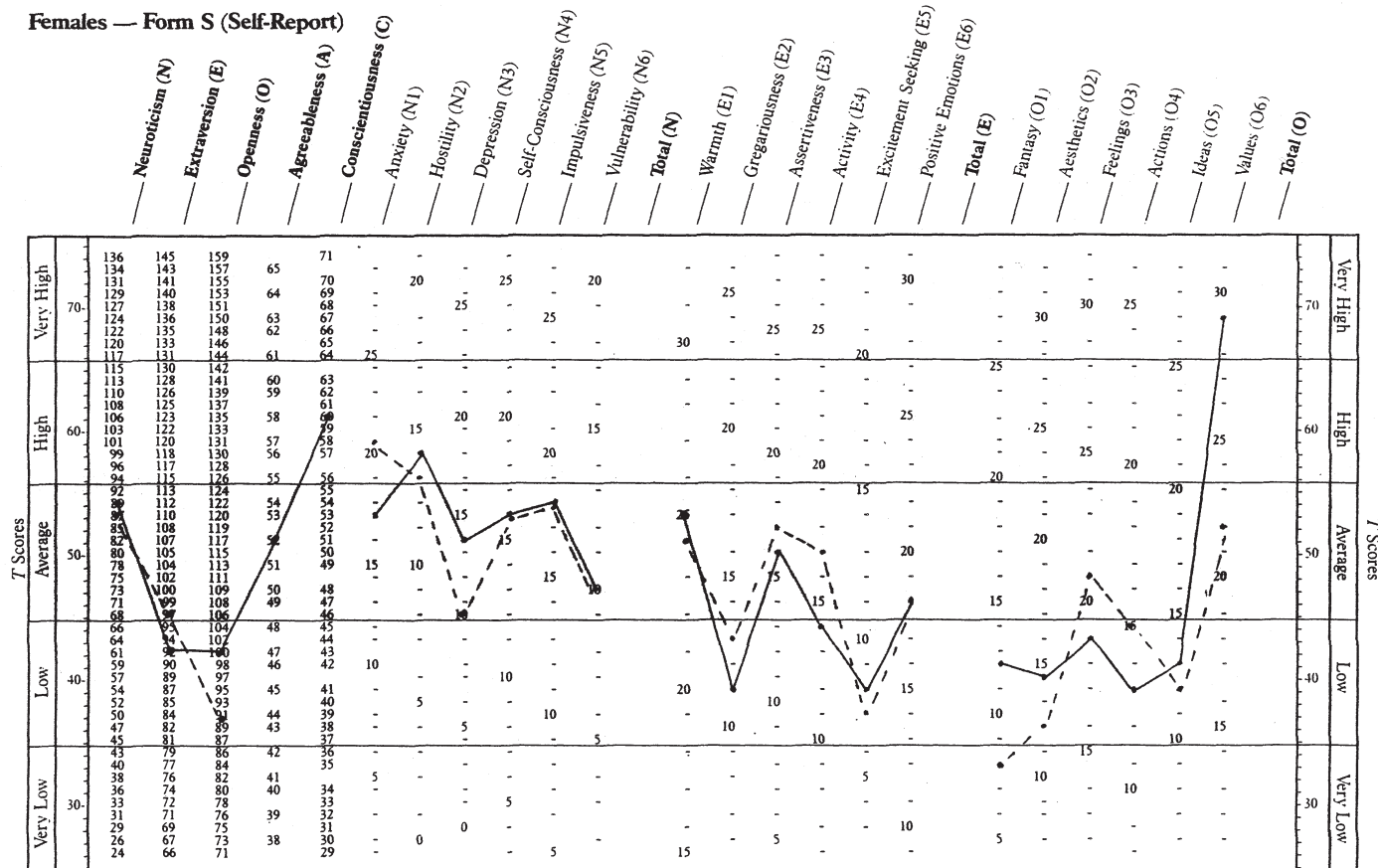


FIGURE 5. Personality profile for Case 1, a 62-year-old homemaker. Profile form reproduced by special permission of the publisher, Psychological Assessment Resources, Inc., 16204 North Florida Ave., Lutz, FL 33549, from the NEO Personality Inventory, by Paul T. Costa, Jr., and Robert R. McCrae. Copyright 1978, 1985, 1989, 1992 by PAR, Inc. Further reproduction is prohibited without permission of PAR, Inc.

Males — Form S (Self-Report)

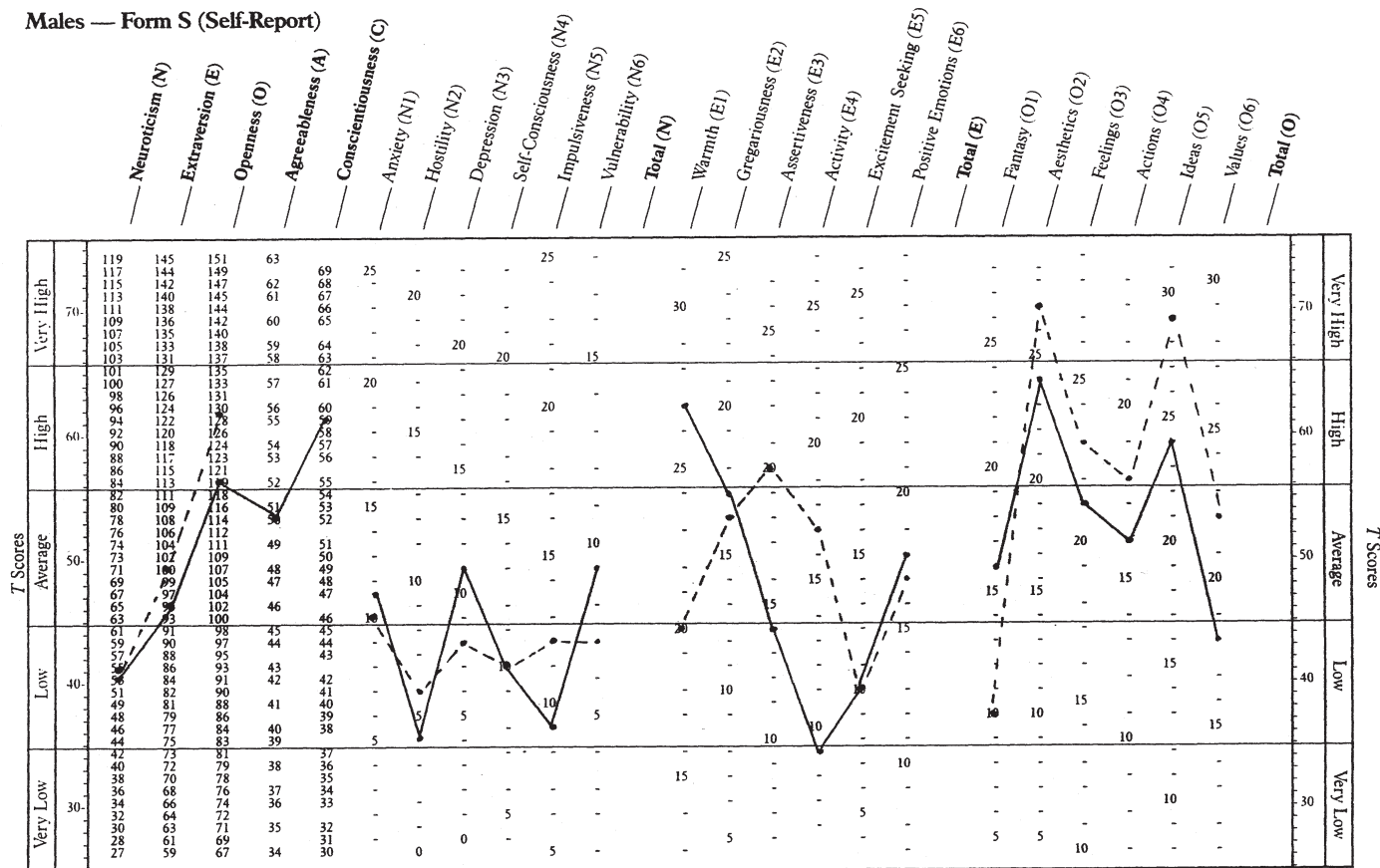


FIGURE 6. Personality profile for Case 2, an 87-year-old minister. Profile form reproduced by special permission of the publisher, Psychological Assessment Resources, Inc., 16204 North Florida Ave., Lutz, FL 33549, from the NEO Personality Inventory, by Paul T. Costa, Jr., and Robert R. McCrae. Copyright 1978, 1985, 1989, 1992 by PAR, Inc. Further reproduction is prohibited without permission of PAR, Inc.

Females — Form S (Self-Report)

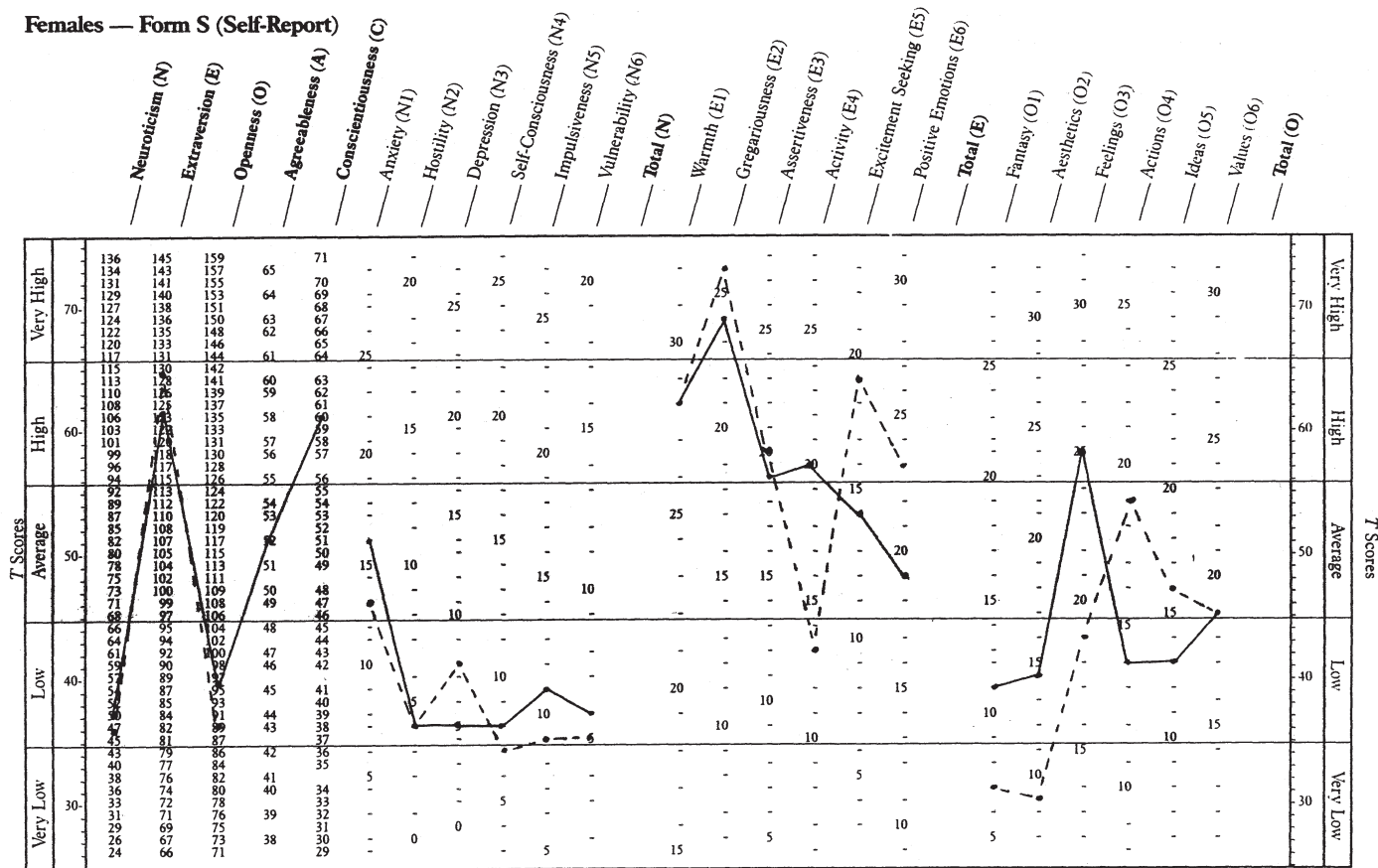


FIGURE 7. Personality profile for Case 3, a 64-year-old social worker. Profile form reproduced by special permission of the publisher, Psychological Assessment Resources, Inc., 16204 North Florida Ave., Lutz, FL 33549, from the NEO Personality Inventory, by Paul T. Costa, Jr., and Robert R. McCrae. Copyright 1978, 1985, 1989, 1992 by PAR, Inc. Further reproduction is prohibited without permission of PAR, Inc.

in other area? Future studies may provide an answer. Overall, her introverted, unadventurous personality seems consistent with her occupation.

Case 2 (Figure 6) is a retired minister and college professor who is well-adjusted, average in Extraversion and Agreeableness, and high in Openness and Conscientiousness. His scores show relatively more change, particularly in facet scales, but stability predominates at the broader level of domain scales. Although he seems to be somewhat lower in several facets of Openness at the second administration, he compensates by higher Openness to Fantasy.

A 64-year-old social worker provides our third case (Figure 7). As we might expect from her profession, we can see that she is a very sociable person, high in Warmth and Gregariousness. Like the professor, she shows some fluctuation for facet scales, but her domain scores are almost identical in 1980 and 1986.

These profiles are typical, showing neither the most stable nor the most unstable of individuals in our sample. Scale scores are not identical on the two occasions, but recall that scale scores are rarely identical even when the interval between tests is only a few weeks. Over the 6-year interval covered here, the resemblance in personality profiles is unmistakable. These people are still the same individuals with the same unique configuration of traits.

IMPLICATIONS: PLANNING FOR THE FUTURE

Imagine the chaos that would result if personality were *not* stable! How could we commit ourselves to marriage if the qualities we loved in our spouse were subject to change at any time? Who would go to the trouble of completing medical school without the faith that he or she would still be interested in medicine years later? How could we vote wisely for politicians if their past diligence and conscientiousness were not a token of their future performance? On what basis would we decide to retire early and move to Florida if we thought that at any moment we might become compulsive workers with boundless energy?

With the exception of those closest to us (whom we may never give up trying to change), we expect people to stay just as they are. Good or bad, we want them to be dependable and predictable so that we can count on them in making plans for our own future. They feel the same

way about us; one of the explanations for stability is the social pressure applied to keep everyone in his or her place. We can shape our own lives, attain our dreams, and find fulfillment for ourselves only if we can make realistic plans for the future. Since so much of life depends on interactions with others, being able to guess how they are likely to respond years from now is essential.

The same is even more true for ourselves. Continuity in personality is a requirement for planning a viable future; it is also a source of the sense of identity. When, in Erikson's last stage, we review our lives and take stock of what we have been and done, we must do so to some extent in terms of our enduring dispositions. If we are high in Conscientiousness, we will take pride in our accomplishments; if we are low, we will recall all the fun we had. Lifelong dreams of writing poetry or fostering a family or dominating an industry make sense only in terms of the basic needs and traits they express. The constancy of personality is a unifying thread that gives meaning to our lives.

* * *

Once we begin to think in terms of stability, it becomes increasingly intuitive. Our parents and grandparents do not seem to change, although our understanding of them may alter greatly as we grow up. History tells us that Beethoven was a rebel at age 20 and at age 50, that Chairman Mao did not grow conservative with age. Hospital and prison records show that tendencies towards mental illness and antisocial behavior are dishearteningly stable. One begins to wonder how the idea of adult development ever arose to begin with.

But are we perhaps missing the whole point? Granted, emotional, interpersonal, and experiential styles may be relatively fixed, but are these the real cores of personality? Or are they instead peripheral characteristics whose stability is no more remarkable or noteworthy than stability in color of eyes or size of feet? Perhaps what is needed is not a different form of personality test or a different personality rater but a wholly different conception of personality in which it is seen not as trait but as process. Perhaps processes change with age.

A Different View

Ego Psychologies and Projective Methods

The finding of stability in discrete traits we have discussed is not likely to impress those who think of personality in terms other than traits. Behaviorists in general do not think of personality at all. They are interested in studying—and especially in controlling—behavior, and traits offer little promise of a way to reshape behavior; in fact, they may represent the greatest resistance to change.

Our concern in this chapter, however, is with a different set of approaches to personality: the psychodynamic psychologies that see personality not as a collection of traits, each operating more or less independently to influence behavior, but as an *organization* of needs, motives, dispositions, habits, and abilities, an organization generally thought to be in the service of certain overarching goals. Depending on the version of psychodynamic theory, these goals might be set by biological instincts, the dictates of culture, or the individual's own experience and convictions.

In this chapter we also turn to an evaluation of projective methods, a form of personality assessment frequently advocated by ego psychologists. Do they make satisfactory substitutes for self-report inventories, or do they provide unique insights? Or does a careful consideration lead to a rejection of projective techniques altogether? And what do they say about stability or change?

DYNAMIC OR EGO PSYCHOLOGIES

To the student of personality theory, it may seem strange to group together the theories of Freud, Jung, Rogers, Murray, Loewinger, Maslow—even Allport—as ego psychologies (in contrast to trait psychologies), but from a certain perspective it is perfectly reasonable. All these writers assume that individuals possess a variety of different and potentially conflicting tendencies, and all assume that the major business of personality theory is to explain how and why all these impulses are channeled and directed into the routine, purposeful, and occasionally irrational behavior we engage in, or the stream of consciousness we experience. Although the term *self* is sometimes used, *ego* is the word most frequently chosen to represent the aspect of the mind or personality that does the organizing.

We cannot hope in a few pages to provide adequate sketches of the theories we will be discussing—dozens of textbooks in personality theory do that already. We would, however, like to try to indicate some of the features they share that set them apart from trait theories. For our purposes we will emphasize the similarities of ego psychologies and minimize the profound differences, which are perhaps better known.

For Sigmund Freud (1933), the ego develops as a mediator among fierce and primitive instinctual impulses (the id), punitive prohibitions internalized in childhood (the superego), and the social realities of the world. It operates through a series of tricks called *defense mechanisms* and is ultimately very much at the mercy of the powers it attempts to control. This division of the psyche is at least as old as Plato, who wrote of appetite, passion, and reason as the elements of human nature, and who held that the ideal was the harmonious balancing of these parts. Freud's framework also seems to parallel the different emphases that have distinguished psychological schools in this century, since it acknowledges the importance of inborn tendencies and environmental pressures, as well as the capacity of cognitive processes—reason and individual choice—to moderate both these influences. In short, it seems likely that every complete psychology will have to deal with these issues in some form; one of Freud's major contributions was to formulate the 20th-century version of an age-old question. Many important personality theorists have made their primary contributions by modifying previous descriptions of the nature or significance of one of the three structures.

Henry Murray (1938) used the same general scheme of dynamic control but changed the cast of characters considerably. Instead of describing primitive impulses, he endowed human beings with a range of needs, some inborn, some acquired; and he proposed that the ego is a strong agent, not a figurehead, in most people, who use reason rather than defense to come up with solutions, compromises, plans, and schedules. Murray also proposed that the superego, the agent of morality within the individual, is not wholly unconscious and immutable. He felt that maturation, or what we would today call *moral development* (Kohlberg, 1971), was not only possible but the general rule.

Gordon Allport, Murray's contemporary at Harvard, developed two theories of personality which he never clearly integrated. He is perhaps best known for his work as a trait psychologist (1966), but he also developed a dynamic theory of personality (1955). He emphasized the purposeful nature of human conduct and the significance of overarching goals, or *proprie strivings*, that guide and organize life.

Abraham Maslow (1954) was also a motivational theorist; his major thesis was that there are broad classes of needs that organize individuals' lives. The history of satisfaction of these needs determines the person's progress through a set sequence of motivations, ranging from the physical through the social to the transcendental needs of the self-actualizing person. In the writings of humanists like Maslow, the element of conflict, central to Freudian thought, is minimized and characterizes maladjusted individuals rather than the natural human condition.

Developmental Sequences

In addition to their similar emphasis on the process of resolving and satisfying needs, most dynamic personality theorists explicitly include theories of human development. Some, like Freud, with his psychosexual stages, or Maslow, with his motivational levels, deal with changes in the nature of the motivational tendencies, the forces that are to be channeled. Others concentrate on growth in the structures that organize the fundamental tendencies. Most of these theorists assert that higher processes such as reason, empathy, and self-control allow mastery of raw impulses in ever more sophisticated ways.

Jane Loevinger (1966) proposed a theory of what she called *ego development*. She distinguished six stages each qualitatively different from the others. The *impulsive* stage is characteristic only of infants and per-

haps of severely disturbed adults, but the other stages are found at all ages. Although each person is supposed to go through the same sequence of stages, many do not outgrow the early ones. In theory, the process is irreversible: Once mature, one can never again think and act in a truly immature way.

The stages Loevinger proposes are perhaps most easily characterized in terms of socialization. Individuals in the early stages are *preconventional*; they want their own way and have few scruples about how they get it. The rules of society are for them merely obstacles to their desires, and although they may follow rules when they know they would be caught otherwise, they have no real allegiance to them. Individuals at the *conventional* stage—where most adults function—have a very different attitude. They believe wholeheartedly in the need for and wisdom of rules, as determined by religion, law, or local standards of taste and etiquette. *Conforming* persons do what they are told without much question; they follow rituals and forms without necessarily understanding them, and they are motivated by the desire to maintain the good opinion of their fellows. *Conscientious* men and women, in the next higher stage, are also loyal to their country and faith but are more thoughtful in their interpretation. They are guided by principles, not customs, and are beholden mostly to their own consciences. At the same time, the principles they follow are dictated by society, and they do not venture to question them. (Note that this usage of “conscientious” is different from ours, but it seems likely that people scoring high on the Conscientiousness factor would also be classified at or above the Conscientious level of ego development. In fact, Einstein & Lanning, 1998, reported an association between ego level and Conscientiousness in college men.)

At the two highest stages of ego development are individuals who can be considered *postconventional*. They have internalized the principles of their social environment, but they have also transcended them. They, too, are guided by principles, but by principles of their own devising; and they are not unwilling to defy social customs that conflict with what they believe is a higher law. Great social reformers and founders of religions are usually at this level, but so are many ordinary individuals who do not make a career out of their ethical opinions.

This description of Loevinger’s stages is, of course, a great oversimplification, in part because it fails to show the importance of emotional

and cognitive changes that accompany the moral development. The form of thought necessary to articulate one's own moral principle is far more sophisticated than that required to say "I want it!" Sensitivity to others' feelings or to one's own unspoken fears or desires is also a far cry from the raw lust or rage that can characterize individuals at the lower ego levels. Progressive levels of ego development, then, are characterized by increasing complexity, differentiation, and integration of thoughts, feelings, and actions.

Dynamic Dispositions

Dynamic theorists often prefer to speak of needs or motives rather than of traits. Motives and needs are dynamic in the sense that they drive behavior, urging us on to do whatever is necessary to achieve certain goals. Some needs, like the need for oxygen, are universal; others, like the need for achievement, characterize only some people or vary in the degree to which they are important to individuals. Motivational psychologists are concerned with the universal ways in which needs are experienced, acquired, and expressed, whereas personality theorists are generally interested in identifying the needs or motives that typify or characterize an individual, usually in contrast to others. Thus, we say that Smith is high in the need for achievement whereas Jones is low in need for achievement but high in the need for affiliation.

If this description of needs sounds familiar, it is because it is very similar to the definition previously offered for traits. We showed in Table 6 (Chapter 3) that the Murray needs measured by Douglas Jackson's PRF can be classified within the FFM. Some theorists, such as J. P. Guilford (1959), regard motives as a class of traits. The only real problem with this conceptualization is that it suggests that there is a clear boundary between motivational traits and other traits. In fact, the distinction between motivational and temperamental traits, or between adaptive and expressive traits (Allport, 1937), is in many respects artificial. Traits as we measure them seem to form syndromes of motives, moods, expressive and adaptive behavior, and attitudes. A hostile man, for example, may be high in the need for aggression, often experience the emotion of anger, have a gruff tone of voice, use direct attack as a way of solving problems, and believe that people will always do their best to cause him trouble; a woman high in openness to aesthetics may

desire to collect art, enjoy beautiful objects, have an elegant handwriting, take lessons to learn painting, and believe the government should support the arts. Traits, then, can be said to have motivational as well as other components. Recall that *disposition* is another word for both *trait* and *motive*. Given the similarities between the two, it seems somewhat silly to decry the “static” nature of traits.

CONTRASTING TRAITS WITH EGO PROCESSES

Trait theories do share with dynamic theories a concern for motivation. But in other respects the two approaches differ, and it is important to consider these differences in judging whether personality changes with age. The most significant differences have to do with the role of ego processes in resolving conflicts and organizing behavior over time.

Conflict and Its Resolution

Dynamic theories of personality tend to emphasize the interaction of the elements they postulate. Freud’s instinctive urges are forever battling against irrational prohibitions and rational restraints for expression. Each bit of behavior we observe and each passing thought in our mind is held to be the end result of this perpetual conflict. At one moment the forces of the ego are in control; the next, a slip of the tongue or pen reveals a momentary breach in the defensive armor. Rarely can any behavior be understood as the expression of any single internal tendency. Instead, drives may merge (as sex and aggression do in sadism), or may be transformed by defense mechanisms into their opposite, or may represent a compromise between the dictates of the superego and the cravings of the id. It hardly comes as a surprise to Freudians when personality traits do not predict behavior very well, since behavior is a function of the constantly changing relative strengths of a host of actively competing forces.

As usual, Freudian theory presents this view in its extreme form, holding that all behavior is defensive and that conflict is inevitable. Later theorists, like Rogers and Maslow, took a less grim view of the human condition, believing that conflict was not inherent in human nature. But many personality theories recognize the need to explain the

mechanisms by which conflicts between impulses or principles are resolved when they do arise.

At each moment we are capable of acting and experiencing in only a very limited way. We are not at liberty to express our assertiveness or indulge our anxiety continually, even if we would like to. We are limited by the opportunities the environment presents, the demands that society and the persons we interact with place on us, and sometimes the competing pressures of incompatible traits: Excitement seeking bids us to drive faster; anxiety urges us to slow down. It is obvious that choices must be made—consciously, unconsciously, or by habit—and it is clear that these choices are not, by and large, random. Sane, mature individuals are not buffeted by internal and external pressures like leaves in the wind. With more or less success, they coordinate their behavior and routinely resolve conflicts. The mechanisms for resolving conflict are an important part of the ego.

For psychoanalysts, conflict is usually resolved through the use of defense mechanisms (A. Freud, 1936/1946). Wishes or memories that would create too much guilt or anxiety because of their conflict with internalized prohibitions are repressed, transformed, or disguised. The individual is not even aware that a conflict exists except when the transformations fail or lead to neurotic symptoms. Whether or not one accepts the accounts offered by analysts, it is certainly true that individuals do learn to control their impulses. Sometimes suppression—deliberate and conscious inhibition—is involved, but often the offending desires never reach consciousness. Children cry when they do not get what they want; adults may feel frustrated, but it rarely occurs to them to cry in such cases.

Murray and Kluckhohn (1953) describe a number of more rational ways to resolve conflicts between tendencies. Using reason, foresight, and self-control, the individual with a strong ego is able to choose between competing courses of action, resist temptations that would lead to undesirable consequences, and resolve the conflicting claims of social pressures, personal preferences, and external necessity. Resolving conflicts need not mean the application of rigid self-control. There are usually opportunities for compromise or for creative solutions. Two needs may be fused (i.e., both may be satisfied by the same behavior, as when a neurotic extravert talks about his imaginary illnesses); or one might be used in the service of another, as when a person closed to values uses her assertiveness to force a conservative position on a group.

Temporal Organization

Another way to resolve conflicts is to allow each tendency expression at a different time. By creating—and sticking to—routines, or *schedules* (as Murray calls them), we can manage to satisfy all or most of our needs on a regular basis. Organizing behavior in time is an essential and distinctive feature of ego processes.

Consider for a moment the course of a typical day. We get up, eat breakfast, go to work, apply ourselves to the necessary tasks, take breaks, talk to our coworkers, eat lunch and perhaps read a book, go back to work, and so on. Trait theorists would be able to make certain predictions about this collection of behaviors: Introverts are more likely to read at lunchtime and less likely to spend a good deal of their day in social conversation; people high in Neuroticism may wake up tired, get angry about the work expected of them, and overeat at lunch; open people are more likely to spend their working hours on tasks like writing, teaching college, or doing research, whereas closed people will more likely be attending to business and following strict routines.

But although information about traits can be extremely useful in predicting specifics of behavior, it cannot by itself account for the structured flow or temporal sequence of behavior. These are ego functions, as important in shaping the life course as in managing daily routines. In addition to schedules, Murray proposed the notion of *serials*, programs of action extended over time and intended to achieve long-range goals. Attending college and law school, working in a law firm, beginning one's own practice, and running for local office may all be steps toward a political career; only with planned and coordinated efforts, pursued diligently over a long period of time, could one achieve this end. Murray holds that the selection of realistic goals and their pursuit over extended portions of the lifespan are major tasks of the ego.

INDIVIDUAL DIFFERENCES IN EGO PROCESSES

Trait theory does not deal with the processes—instrumental learning, formal reasoning, creative planning, defensive distortion—by which behavior and experience are molded on a moment-to-moment basis. But although ego psychologists make these processes central to their

theories, most are also concerned with characterizing individuals. People differ in the needs, values, and goals that must be organized and harmonized by the ego, but they also differ in the ways in which the ego goes about its task of coordination. It is not simply that people make plans for the future: Some plan wisely, some poorly, some hardly at all. Individual differences in ego processes are an important issue in personality theory.

Jane Loevinger's theory of ego stages is a clear example. Men and women in the lower stages of development are subject to strong and primitive impulses and tend to express them immediately. Because their ego functioning is undeveloped, they cannot resolve conflicts effectively or set up efficient schedules. In the long run they are unable to satisfy many needs, at least without provoking guilt or reprisals from others. Persons with higher levels of ego development have far more sensitive and sophisticated ways of integrating their needs and values. On the other hand, the lives of individuals with high ego levels are correspondingly more complex, and they are troubled by abstract problems to which persons with low ego levels are oblivious. Thus, high ego development does not necessarily mean greater satisfaction or happiness (McCrae & Costa, 1983b).

Other theories of individual differences in ego functioning are attentive to overall adjustment. *Ego strength* is a characteristic discussed by a number of psychoanalytic writers, who generally define it as the ability of the individual to manage successfully the competing forces of instinctual demands, internal prohibitions, and social reality. Henry Murray used the term to describe characteristics needed to lead a well-ordered life. People thought to be low in ego strength are less intelligent, more ethnocentric, and generally more prone to many kinds of psychopathology (Barron, 1980).

Jack Block (1965) distinguished two dimensions of ego functioning. *Ego resiliency* is defined as the individual's ability to adapt to new demands. Individuals high in ego resiliency are resourceful and flexible; those who are low are touchy, moody, and uncomfortable in their world. *Ego control*, a second dimension, is seen in characteristic levels of impulse control. Overcontrolled individuals are overconforming, narrow in interests, and interpersonally distant; undercontrolled people are spontaneous, inclined toward immediate gratification of impulses, and willing to act in new and untried ways.

Traits and Metatraits?

Concepts such as ego level, ego strength, and ego control are intended to describe the functioning of the ego in different individuals, but as enduring characteristics they resemble traits much more than ego processes. One way to distinguish them from more mundane traits is to grant them the status of superordinate traits, or *metatraits*. These are the enduring dispositions that govern the expression and integration of other dispositions. Loevinger (1966) came close to this formulation when she called ego level the “master trait.”

But as we begin to examine the idea of metatraits closely, their distinctiveness begins to evaporate. Block’s description of ego resiliency (1965; see also J. H. Block & J. Block, 1980) makes it sound very much like low Neuroticism, and a measure of ego resiliency based on MMPI items showed a strong negative correlation with NEO-PI Neuroticism ($r = -.70$, $N = 274$, $p < .001$) in the BLSA. Similarly, ego control seems to represent low levels of Impulsiveness, Excitement Seeking, and Openness, and these hypotheses are also confirmed (r ’s = $-.27$, $-.54$, and $-.34$, $p < .001$). Frank Barron’s ego-strength scale (1980) is known to correlate with measures of maladjustment and Neuroticism. And Loevinger’s master trait of ego level (1966) is strongly related to intelligence as well as to personality (McCrae & Costa, 1980). A study we conducted in Boston employed the sentence completions of 240 men to examine ego levels using Loevinger’s own instrument, the Washington University Sentence Completion Test. The scores we obtained were correlated with our measures of personality and intelligence. We found, as she has repeatedly done, that there is a marked correlation with IQ (about .50), showing that higher ego levels are more often found in more intelligent people. But we also found that there was a significant correlation with Openness to Experience, even when we controlled for intelligence. More open people are rated as being more mature in the Loevinger system.

One interpretation of these conceptual and empirical correspondences would suggest that traits and metatraits are not really distinct. The same traits that are organized by the ego are responsible for the individual differences in the way the ego organizes them. Intelligence is a trait, but it also modifies the ways other traits are expressed: Introverts read books; intelligent introverts read difficult books. Again, anxiety is a trait, but the disposition to experience anxiety may disrupt the efficient

functioning of other traits. Test anxiety, for example, may lead to poor performance despite high intelligence.

If these speculations are correct, we do not need to introduce new traits to explain differences in ego organization or the resulting differences in the life structure. Instead, we must recognize that in order to describe fully the significance of a trait, we must not only specify the behaviors and feelings associated with it but also identify the effects the trait has on the operation of other traits.

The other side of this interpretation is that in inferring traits we ought to look not simply at specific behaviors but at the whole pattern of the individual's life. And this is precisely what we advocated in Chapter 2 as the best basis for assessing traits.

DEVELOPMENTAL CHANGES IN EGO PROCESSES

If the essence of personality is the organization and integration of experience and behavior, the most important question about age and personality would be what happens to the major ego functions with age. It is in this context, perhaps, that the issue of growth or decline in personality makes most sense. After all, what do we mean by growth in personality? It is easy to see that becoming taller or more intelligent is growth, but how does personality grow? Although we assumed in Chapter 4 that any change in the average level of a trait could be interpreted as either growth or decline, in fact it would have been difficult to interpret an increase in introversion or Openness to aesthetic experiences in these terms. Not all changes represent growth or decline.

But we could probably make a good case for the contention that there is a natural direction of growth in ego functioning. Loevinger suggests as much in her developmental stages, as does Erikson (1950) in his stages of psychosocial development. Many other writers have also proposed criteria of psychological maturity that could properly be seen as growth in personality and would generally be considered attributes of ego functioning. Bühler and colleagues mention formulation of a stable identity, willingness to assume responsibility, ability to form significant and lasting attachments to others, self-acceptance, emotional security, and commitment to goals (Bühler, Keith-Spiegel, & Thomas, 1973).

Most of us would agree that these are signs of maturity, and we would probably suspect that adults possess more of them than do ado-

lescents. But does maturity increase with age once we have become adults? Or do older individuals regress to more primitive ways of resolving conflicts, dealing with others, and organizing their activities?

If the argument we developed above is correct—if ego characteristics are determined by such traits as anxiety, depression, warmth, and openness to ideas—it would follow that there ought not to be major changes in the functioning of the ego across the adult lifespan, since the traits that determine it are themselves stable.

Consider the case of ego development. If individuals continue to mature in the period of adulthood, the ego levels of older individuals should be higher than those of younger individuals. But we know that ego level is related to the traits of intelligence and Openness and that both of these are generally stable in adulthood. Trait psychologists would predict, then, that there should be no changes in ego level. And cross-sectional data from the Boston sample clearly confirm that prediction (McCrae & Costa, 1980), as do other independent studies (Vaillant & McCullough, 1987). Longitudinal studies have shown that the predicted changes in ego development are seen as individuals move from childhood to adolescence, but there is no evidence of change after adulthood is reached. If ego level is indeed the master trait, this is one more evidence of stability at all levels of personality.

Conclusions on stability or change in ego processes are perhaps premature. It is much harder to describe the organization of elements than the elements themselves. Additionally, there is no consensus on what properties of the ego we should be measuring: Is control of impulses the key, or is it the arrangement of attitudes and beliefs? Should we be concerned with the normal daily functioning of individuals, with their responses to crisis situations, or with their long-range plans? Or must we develop a battery of tests to consider each of these separately?

And what kind of tests? It may be that individuals, although able to report adequately on emotional reactions or vocational preferences, are incapable of judging the more profound styles that tie together all these discrete elements. Perhaps an analysis of a deeper level of personality is needed. Projective tests are supposed to give a picture of the deeper layers of personality, and we will shortly review what they have to tell us about age and personality. But just as we were compelled to describe the rationale and evaluate the efficacy of self-report measures in Chapter 3, so now we must do the same to lay the groundwork for an evaluation of projective tests.

PROJECTIVE ASSESSMENTS OF PERSONALITY

For reasons partly historical and partly theoretical, psychologists of the dynamic school have usually preferred evaluations of personality based on something other than self-reports. The first and most common argument offered against self-report assessments is that individuals may not be aware of important aspects of their personality and may in fact systematically distort their perceptions and reports of themselves. This is a perfectly reasonable point of view for Freudians, who locate the variables of real interest in the unconscious. We cannot expect individuals to report accurately the ways in which they deceive themselves and try to deceive others; yet, according to psychoanalysts, these defenses are among the most important aspects of personality.

As a result, clinicians and researchers have devised a number of methods of assessing personality indirectly. Most of these tests are called *projective*, from the belief that the individual unwittingly projects his or her personality into the task. Asked to describe a picture, individuals probably imagine that their responses tell about what they see. The psychologist, however, believes that the responses tell about the subject. A variety of techniques have been used to outsmart the subject in this way. The individuals being assessed may be asked to give associations to words, report what they see in inkblots, make up stories to accompany the cards of the Thematic Apperception Test (TAT), draw pictures, or write completions for unfinished sentences.

The scoring of these materials is, of course, very different from summing up the items on a scale. A judge (usually human, but occasionally a computer) must consider the subject's production in the light of some set of rules developed to judge the characteristic of interest. Sometimes the rules are quite specific; sometimes they require extensive intuition. The scoring system may be based on the empirical strategy of contrasting the responses of one group with those of another, or it may be derived from purely theoretical considerations. Psychoanalytically oriented researchers often propose scoring systems based on the presumed symbolic significance of responses. There is a strong tradition of assessing needs through the TAT on the assumption that the concerns of the story's hero are those of the story's creator (McClelland, 1980; Murray, 1938). A story about long struggles leading to a successful career, for instance, might be interpreted as evidence of a high need for achievement.

Some psychologists prefer unstructured projective methods to self-report questionnaires for very different reasons. McClelland (1980) distinguished between what he called *operant* and *respondent* tests. The former consist of materials—a picture, a blank sheet of paper, a sentence fragment—that the subject can organize and interpret for him- or herself; the pictures on TAT cards are the primary example. Respondent tests are exemplified by questionnaire items in which the subject must respond within a strictly limited number of alternative ways to a fixed set of items. McClelland objected that this regimentation, which makes scoring of respondent tests so convenient, introduces artificiality. An older individual, for example, might find all the items in a test designed for college students irrelevant to her, though if urged she would still fill it out. What would the responses mean in that case? Operant tests are not subject to this criticism. The subject can always make a picture relevant by choosing what elements to respond to, or by creating a story that goes beyond the picture toward whatever concerns he or she has. Similarly, McGuire (1984) argued that psychologists should study the spontaneous self-concept, as revealed in free descriptions, to learn how individuals really see themselves.

Jane Loevinger made a different point in defending her choice of a sentence completion test to measure ego development. What is of primary interest, she maintained, is not so much what the person decides as how he or she makes the decision. Asking the subject to complete stems (e.g., “What I like best about being a man is . . .”) gives the researcher an opportunity to observe how the subject will structure and organize his responses, free from the prompting of the investigator. Since it is the processing of the ego that Loevinger wants to measure and since most situations in life present the same ambiguity as pictures or half-formed sentences, it can be argued that tests that simulate these conditions should give the best picture of the ego at work. In effect, such tests provide samples of ego functioning, rather than reports about how the individual thinks he or she is.

There are thus three major lines of argument about the superiority of projective tests over questionnaires: The first holds that the real personality is hidden from the subject and can only be inferred indirectly from signs and symbols that the trained psychologist must interpret. The second claims that structured questionnaires are often irrelevant and only unstructured material can convey an accurate representation of the individual's nature. The third argument supposes that personality is best assessed by observation of its workings in an unstructured situa-

tion. All would concur in the belief that answers to personality inventories are largely irrelevant.

Problems in Projective Methods

Historically, these arguments have been very powerful. Projective methods continue to be among the most widely used of psychological tests in practice and in research (Lubin, Larsen, Matarazzo, & Seever, 1985). For many years, however, psychometricians and those who have reviewed the empirical literature have repeatedly pointed to serious problems in projective methods, most recently in a monograph published by the American Psychological Society (Lilienfeld, Wood, & Garb, 2000). Our examination of the issue has led us to side with the critics, but the reader should realize that this is a controversial, sometimes unpopular position to take.

What kinds of problems do projective and other unstructured tests pose? They are highly inferential, ignore other factors that contribute to responses, are unreliable over time, and are inadequate as samples of behavior. The bottom line is that they frequently fail to show evidence of external validity; that is, they do not predict the kinds of outcomes they are supposed to.

Clinicians in the psychoanalytic tradition are trained to seek out the hidden meanings behind such seemingly innocuous or meaningless phenomena as dreams, slips of the tongue, or word associations. The Rorschach inkblots were designed to provide standard ambiguous stimuli onto which the patient could project his or her inner conflicts and fixations, and clinicians were supposed to interpret these responses in terms of dynamic theories of symbolic expression. A skilled interpreter can often create a persuasive reading of inkblot responses, but it is usually impossible to tell if the interpretation is correct. We certainly would not want to trust peer or spouse ratings in such a case, since few friends or relatives are competent to judge the degree of oral fixation or intrusion of the anima. The most promising test of the quality of an interpretation is its agreement with the interpretations of other qualified clinicians; yet extensive research has typically yielded little evidence of agreement on inferences about the so-called deep level of personality. Mischel (1968) put it this way:

Statements about "deep-seated unconscious anxieties," "latent homosexual tendencies," "reaction formations," and other similar inferences about

inferred covert dynamic processes or structures . . . may provide intriguing reading. No matter how fascinating personality descriptions may seem, however, they have little value when competent judges cannot agree about them. (p. 121)

Interpretation is difficult in part because so many things other than personality influence the responses. When we ask subjects to draw a picture or tell a story, their artistic abilities are likely to be much more in evidence than their intrapsychic dynamics. Under the naïve assumption that the hero of the story represents the personality of the storyteller, we would have to infer that novelists, who are capable of describing a huge range of human characters, have multiple personalities. Perceptual problems (which are common in the elderly), verbal skill, and motivation all have a much greater impact on projective responses than they do on objective test results. Indeed, projective tests may be most valuable as samples of cognitive behavior—not personality.

Reliability has always been a thorny issue for projective tests. Personality is supposed to be consistent across time and situations; it is, almost by definition, that which endures as the momentary pressures of the environment and the body vary (Maddi, 1980). The specific behaviors may well change, but the psychologist should be able to infer the underlying unity throughout the changes. It seems reasonable to expect that the projective test results, which allegedly mirror personality, would show the same degree of consistency across different forms of the test or at different times. This is the criterion of reliability, and it has always proven to be a sore spot in projective testing. TAT stories change from one time to another, and so do the scores derived from them (Winter & Stewart, 1977), perhaps because needs change as a result of their expression on the test (Atkinson, Bongort, & Price, 1977).

But if the expression of needs is subject to such marked fluctuations, we have no way of knowing whether the responses we obtain on the first test are characteristic of the individual or simply a reflection of the immediate situation. The TAT measures may be accurate reflections of the motivational *state* of the individual, but if they are not reliable over time, they are poor indicators of the *characteristics* of the individual. The same argument, of course, also applies to projective tests of processes other than needs: defenses, complexes, attitudes, and cognitive styles. Demonstrable retest reliability is essential to the measure-

ment of individual characteristics, and many projective tests fail to meet even minimal standards of reliability.

There are significant problems with the idea that unstructured tests can be used to obtain an interpretable sample of the operations of personality. A questionnaire item that asks "Are you often irritable and angry?" requires the subject to review his or her behavior over a period of months and to estimate the typical level of this state. To the extent that the person is able to weigh accurately his or her recollections, this item gives an average based on thousands of hours of living in a wide range of situations. By contrast, observing the responses of the individual to an inkblot or sentence fragment gives only a tiny sample of behaviors, all within the thoroughly artificial situation of a psychological testing session. One of the recent conclusions from the trait-consistency controversy is that any single instance of behavior is likely to be a very poor indicator of the average level (Epstein, 1979); if we really want samples of ego functioning, we must plan to spend a great deal more time and effort than we currently do.

Finally, and most compellingly, the results of careful research studies using projective tests to measure traits or predict outcomes have been almost uniformly disappointing. In a detailed review of studies on projective tests conducted between 1950 and 1965, Suinn and Oskamp (1969) concluded, "There are only a few things which the clinician can predict from these personality tests with confidence that his judgments are being made on a scientific basis. The remainder of his predictive work is still based on faith or on theory rather than on evidence" (p. 117). A review based on only five studies (Parker, Hanson, & Hunsley, 1988) came to a more sanguine conclusion on the validity of Rorschach measures, and they continue to have their defenders (Meyer, 2000), but it is safe to say the weight of evidence is still against the scientific value of most projective measures (Lilienfeld, Wood, & Garb, 2000).

Conscious versus Unconscious Elements

One of the most intriguing aspects of Suinn and Oskamp's review (1969) is the consistent finding that the most nearly valid tests were those that agreed most closely with self-report. Sentence completion tests, for example, that come close to direct self-reports are among the better-validated projective tests. Rotter and Rafferty (1950) constructed a sentence completion test of maladjustment that was significantly cor-

related with both clinical judgments and self-reports of adjustment. In our study of ego level and Openness to Experience (McCrae & Costa, 1980), we examined responses to Loevinger's sentence completion stems for evidence of Openness (without knowledge of how subjects had rated themselves). We judged subjects to be *open* when we found statements showing flexible views of rules, rejection of traditional sex roles, intrinsic interest in experience, or playfulness. When we compared our judgments with the self-reports of Openness the same subjects had given us, we found agreement in 78% of the cases.

But projective assessments that make inferences about unconscious processes are more likely to fail to agree with self-reports. In these cases, there is usually also little agreement with any other external criterion of importance, such as suicide, sexual orientation, or improvement in therapy. Fascinating as the symbolic interpretations of the Freudians are, there is simply no good evidence that they are correct.

What all of this suggests to us is that the unconscious level of personality either cannot be measured by projective techniques or does not have much influence on human conduct. No one who has read Freud or Jung could fail to appreciate the marvelous insight that there are levels of the mind that operate by a logic different from that of our waking consciousness; that these are more characteristic of children than of adults; that they appear in dreams, in artistic productions, in psychotic delusions, and in culturally shared myths; and that the primitive ways of thinking we have inherited from our evolutionary ancestors often appear in them in undisguised form. The data collected over the past 70 years, however, have convinced us that these curious relics of our past have relatively little to do with our daily life or with such important outcomes as happiness, response to stress, or political beliefs.

PROJECTIVE TESTING AND THE STABILITY OF PERSONALITY

Let us summarize the argument so far and then proceed to a more specific consideration of the relationship between projectively assessed personality and aging. We have considered the reasons for preferring projective tests to self-reports and have rejected them. Some projective tests aim at a deeper level of personality, but that deeper level, if it is indeed tapped by such tests, does not appear to have much influence on

important aspects of personality, at least in psychiatrically normal adults. Some unstructured tests aim to sample the ego, to judge it by its spontaneous functioning in an unstructured situation. But there is strong reason to believe that the samples obtained are generally inadequate and that when they approach adequacy they simply duplicate information that could more easily have been obtained by direct questioning.

We might, therefore, dismiss projective testing altogether and decline to review a literature based on such faulty instruments. We have noted that sentence completion tests, the most reliable of projectives, show no cross-sectional relationship of ego development to age in adulthood (McCrae & Costa, 1980). At the same time, it will be worthwhile to consider some examples of aging and personality research using projective techniques, if only as a way of illustrating the points already made.

Inkblot Tests

Cross-sectional studies of inkblot responses do in fact show a regular progression of age-related changes. Old people make fewer and more concrete descriptions of what they see and tend to give global responses instead of picking out minor details (Kahana, 1978). Both cross-sectional and longitudinal studies (Ames, 1965) have verified this tendency; but before concluding that here at last is evidence of personality change, we must bear in mind that the Rorschach Inkblot test is used to assess cognition and perception as well as personality. There is no doubt that there are age-related changes in learning, memory, perceptual closure, and visual acuity, and these, rather than personality change, may well account for the age differences. Caldwell made that argument as long ago as 1954, and Eisdorfer (1963) provided some strong evidence in favor of it. When he matched old and young subjects on cognitive abilities as measured by the Wechsler Adult Intelligence Scales, or WAIS, he found no differences in Rorschach responses.

We (Costa & McCrae, 1986a) conducted a 2-year longitudinal study of age differences and changes on scales of the Holtzman Inkblot Technique or HIT (Holtzman, 1961). The HIT is intended to provide a psychometrically superior version of the better-known Rorschach test. Respondents are shown 45 blots and asked to tell what each looks like. Responses are scored for 22 variables, ranging from reaction time (how

long it takes for the subject to offer an interpretation) to the content of the response (animal, human, anatomy) to the definiteness and appropriateness of the response for a given blot. These scales, alone and in combination, can be given psychodynamic interpretations (Hill, 1972).

We originally administered the HIT to 93 men and women in the BLSA aged 25–90. When correlated with age, only four of the 22 HIT variables showed significant associations, and all of these were small. Further, when the test was readministered to 44 of the subjects about 2 years later, only one of the four cross-sectional findings was replicated by longitudinal changes: Scores on the HIT Hostility scale were higher in older subjects and also increased over time, suggesting a possible maturational effect. However, this may well have been a chance finding, since Overall and Gorham (1972) did not replicate it in a much larger cross-sectional study.

We also examined stability coefficients for the 44 subjects with two administrations. These correlations ranged from .07 for Form Appropriateness to .73 for Form Definiteness; most of the HIT variables gave evidence of moderate to substantial stability over the 2-year interval. Finally, we correlated the HIT scales with measures of Neuroticism, Extraversion, and Openness, but we found no consistent pattern of associations. In particular, the HIT Hostility score was unrelated to Neuroticism. (Later analyses showed that none of the HIT scales was related to Agreeableness or Conscientiousness either.)

Interpreting these results is difficult. Because they are not correlated with measures of the five factors, we cannot group the HIT scales into familiar domains of personality—in fact, we could conclude that the HIT does not measure personality at all. Perhaps it measures some form of perceptual or cognitive style, or perhaps merely idiosyncratic preferences: When looking at inkblots, some people tend to see animals, some to see people. Whatever it is that the HIT scales measure appears to be relatively stable over a 2-year period and perhaps for much longer. We found little evidence of age-related change in HIT scores. If the HIT measures unconscious layers of personality, they do not appear to change with age.

TAT Studies

The Thematic Apperception Test (TAT) was introduced by Murray (1938) as a way of inferring motives or needs of which the individual

might not be consciously aware. Veroff, Depner, Kulka, and Douvan (1980) used the TAT for this purpose in two national surveys, conducted in 1957 and 1976. They scored the stories for evidence of the needs for achievement, affiliation, and two forms of power: fear of weakness and hope of power. When controlled for time of testing and education, men showed higher hope of power in middle age and women showed lower need for affiliation in old age. In each sex, the other three needs showed no significant cross-sectional differences.

The technique of eliciting stories from pictures as a basis for assessing needs, traits, and defenses has been particularly popular with gerontologists. Gutmann (1970) used TAT cards for his research on the concept of ego mastery. He scored TAT stories in terms of active, passive, and magical mastery. Cross-cultural longitudinal studies suggested that as men age they pass from a stage of active mastery (in which TAT heroes forcefully tackle the problems of their world) to a stage of passive mastery (in which they acquiesce in the demands of the environment) and then to a stage of magical mastery (in which problems are denied or wished away). Gutmann's description of ego mastery style as a measure of coping style or adaptational ability suggests that it should have some consequences for the ways in which older people adapt to stress.

The use of magical mastery in old age is consistent with some clinical views that hold that older people tend to use more primitive forms of defense, such as blatant denial of reality (Pfeiffer, 1977). These views may accurately reflect the behavior of older people with dementing disorders, but they are highly suspect when applied to cognitively intact older individuals. Once we move away from the patently bizarre delusions of psychotic or demented individuals, it is often difficult to know if behavior or attitudes are truly reflections of denial (Costa, Zonderman, & McCrae, 1991). For example, older individuals, despite the proximity of death, generally do not show a heightened fear of death (Kastenbaum & Costa, 1977). To claim this as an instance of denial, we would have to assert that death really is a threat that should be feared—and we have no objective basis for such an assertion. Some people view death as a release or as a passage to a new and better life. Are they merely rationalizing? Who can say?

Stereotypes of old age have portrayed it as a dismal portion of the lifespan, and evidence that older men and women are generally as happy and satisfied with life as are younger individuals (Costa & Mc-

Crae, 1984) is taken by many as evidence of defensive distortion. But the appearance of happiness among older people is so common that we would have to assume that older people universally resorted to the use of primitive defenses, and this would be inconsistent with the fact that most cope relatively well and realistically with events such as retirement, death of a spouse, and personal illness.

Objective studies of age and coping show no tendency for older people to use magical denial in solving real-life problems (McCrae, 1982a). In two studies, BLSA subjects were asked to describe how they had coped with recent stresses, chosen either by us or by the subjects themselves. Using a self-report checklist of ways of coping, we assigned them scores on 28 different coping mechanisms, including rational action, distraction, denial of affect, intellectual denial, and passivity. We found age differences in the kinds of stress older people faced—generally more threats to health and fewer occupational and family challenges—but when the responses had been corrected for these differences only two consistent findings emerged: Younger men and women were more likely than middle-aged and older ones to indulge in escapist fantasy and hostile reactions. There was no evidence whatsoever that magical solutions were chosen by the elderly.

Gutmann himself is careful to point out that the progression he posits occurs on the level of fantasy, not of reality. Among the Highland Druze of Israel, in particular, the old men are respected as leaders and decisions makers; they are anything but passive (Gutmann, 1974).

So it seems that neither self-reports nor direct observations of behavior show parallels to the developmental sequence inferred by Gutmann from TAT responses. Perhaps these changes result from cognitive or perceptual shifts, or perhaps they are changes in a deeper level of personality. In any case, they appear to have little impact on social or emotional functioning.

A few longitudinal studies have used the TAT to assess individual stability of projectively assessed dispositions. Skolnick (1966) reported 20-year stability coefficients ranging from .21 to .34 for the needs for power, affiliation, aggression, and achievement, although the findings were not consistently replicated across sexes. Britton and Britton (1972) also reported significant stability coefficients for TAT measures of personal and social adjustment over intervals of 3, 6, and 9 years in small samples of men and women aged 65–85 at first testing. The magnitude of these correlations is far smaller than the .60–.80 that we are accus-

tomed to seeing from self-reports, but that is to be expected. Relative to the limited reliability of projective tests, the modest long-term retest correlations from TAT studies can be seen as further evidence of stability.

Age and the Spontaneous Self-Concept

We have argued that unstructured methods that require little inference on the part of the researcher and thus approximate self-reports have generally been shown to be more valid than other projective methods. Measures of the spontaneous self-concept may combine the best of both worlds. In these techniques, individuals are simply asked to describe themselves, either in a few paragraphs or as a series of responses to the question, "Who am I?" (Kuhn & McPartland, 1954). These are certainly self-reports, but they are free of the usual constraints of questionnaires that ask a series of specific questions.

We studied the spontaneous self-concept of men and women in the BLSA (McCrae & Costa, 1988a). Subjects provided 20 different responses to the question, "Who am I?"; two raters categorized their responses. Recall that the self-concept is composed of many kinds of elements, of which personality dispositions are only one. Most people described themselves in terms of social roles ("a loving mother," "a retired teacher"), physical characteristics such as health or attractiveness, and activities and attitudes ("I enjoy boating and camping"). About one-quarter of the responses specified personality traits ("I am hard-working and productive"). In addition to categorizing each separate response, our raters gave overall impressions of the individual's standing on each of the five personality factors and on the general level of self-esteem.

There were several findings of note from this study. First, the impressions of personality that raters drew from the patterns of responses were fairly accurate: The two raters generally agreed with each other, and in most cases their combined ratings were significantly correlated with both self-reports and peer ratings on the NEO-PI. We apparently get much the same information when we ask specific questions of subjects or when we ask them for a general description of themselves and then evaluate the specific responses they make in terms of our basic dimensions. Second, there was little evidence that personality expressed in the spontaneous self-concept was influenced much by age. Older

subjects were somewhat less likely to describe themselves in terms of traits that reflected emotional instability or Neuroticism, but none of the other personality factors was significantly associated with age. Self-esteem was also unrelated to age; instead, it was associated with low levels of Neuroticism and high levels of Extraversion. Adjusted extraverts have high self-esteem at any age. Third, age was only modestly related to the other scoring categories. Older subjects were somewhat less likely to describe themselves in terms of family relationships and somewhat more likely to mention daily activities and attitudes. Age itself was rarely mentioned as a salient part of the self-concept. Older people may think of themselves as citizens, parents, curious individuals, or gardening enthusiasts, but they apparently don't think much about being old. Perhaps they have learned that age itself is not very important for describing what they are like as persons.

* * *

This chapter set out to consider criticisms of the stability position that might be offered by personologists with different theoretical and methodological preferences. We conceded that the processes by which behavior is integrated and organized, both moment by moment and over the sweep of a lifetime, are not explicitly addressed by trait models, as they should be in a complete theory of personality. But we also argued that individual differences in these processes, including differences between young and old people, were likely to have trait-like characteristics and that, in fact, the traits we normally measure are as likely to influence the organization and integration of other traits as they are to influence specific behaviors. To the extent that this is true, we might argue that there are probably no age differences in ego processes.

But is that true? Or are there other organizing features of personality that have eluded the grasp of self-report trait measures? To answer this, we turned to the favored resort of theorists who concentrate on ego processes—projective methods. We found that, to the extent that they depart from the information that could be obtained from self-reports, these measures are generally unreliable and invalid, except perhaps as samples of cognitive behavior or as evidence of shadowy processes (like mastery style) that have little relevance to daily functioning or the course of life. To the extent that they mirror conscious concerns and

tendencies, as in measures of the spontaneous self-concept, they tend to give evidence of stability in personality.

At this point the astute reader will have noticed that the issue has not really been resolved. We admitted that self-reports might not adequately assess the functioning of the ego and then suggested that we examine projective methods. But our assessment of projective techniques hardly leaves us with confidence in any testimony they might offer. Is it not possible, you may ask, that there really are profound changes in the fundamental structure of personality but that projective tests, with all their limitations, are simply inadequate to the task of finding and documenting them? And if we cannot trust self-reports or projective tests, is there another way to explore the deeper levels of personality? Is there a way to sample the ego at work without letting the ego do the sampling? Is there a way to look for age-related differences in the organization of behavior independent of the trait conceptions embedded in our measures? Perhaps an objective and trained observer, sensitized to issues of personality organization and allowed to collect whatever information seemed relevant, could accomplish the task. Perhaps what we need is an intensive clinical interview. Many people have thought so, and some of the most important theorizing on adult development has been generated by this method. We examine it next.

Adult Development as Seen through the Personal Interview

The interview is a venerable psychological technique with many variants: The clinical interview is used to gain diagnostic information; the therapeutic interview, to help the client think through his or her problems; Daniel Levinson's "biographical interview," to gain a sense of the individual's life; the research interview, to gather data. Some interviews, such as those conducted by survey researchers and pollsters, are really questionnaires administered verbally rather than in writing; of these we have little to say, since they are essentially self-report instruments. But most interviews involve freer questioning at the discretion of the interviewer, and most require judgments by a rater (who may or may not be the interviewer) rather than simple tabulations of answers. It is these sorts of interviews that we need to consider next, since they have played an important role in theories of adult development.

As a technique for gathering information on personality, the interview has a privileged status. Clinicians, who have always dominated personality research and theory, are accustomed to talking directly to patients and to making diagnostic judgments about their condition; it is understandable that they would put particular faith in the results of interviews. The same principles and practices that are used in therapeutic interviews can be applied to normal subjects of pure research.

According to its proponents, the interview has all the advantages of self-report instruments and more. The same questions can be asked by either method, so that the unique experiences of the individuals and their wide knowledge of themselves can be exploited. But in the interview, these self-reports need not be taken at face value.

Perhaps the respondent does not seem to understand the question—then it can be rephrased. Perhaps the answers are inconsistent—then the inconsistency can be pointed out or the issue probed more subtly. Perhaps the subject is squirming too much when certain topics are mentioned. The canny interviewer, who is normally a trained professional, can use all these clues to qualify and evaluate the self-reported information. For many researchers the professional interview is the gold standard by which other personality assessments are judged. Farrell and Rosenberg (1981), for example, administered self-report questionnaires to 500 subjects drawn to represent the population in general. But most of the conclusions in the book they wrote on their studies are based on interviews with 20 of these men (and their families). The authors simply did not trust the self-reports to give them the complete and unvarnished truth. (As it turned out, the interviews upheld the major conclusions of the self-report study, although many more complexities were uncovered.)

Before deferring completely to the superiority of the interview, however, we should weigh some of its limitations. The results of each interview depend on human judgment, and even expert judgment is fallible. Interviews are typically conducted in 1 or 2 hours; at best they may last 8–10 hours. The interviewer or the rater who interprets the interview must formulate opinions about the person and his or her life on that rather slim basis. What if the subject is nervous and unused to talking about himself? What if she is having an off day? What if the interviewer simply misreads them, thinking that his jokes are intended seriously or that her serious comments are jokes? Is it really wise to try to second-guess the meaning of an individual's statements?

Even more unsettling is the fact that interview methods are subject to certain kinds of error that are more serious because they are more systematic. If a hundred 40-year-old men are asked how they feel, some will certainly exaggerate how bad things are, some will surely try to cover up their distress; some will be having unusually bad days, some unusually good ones. None of the self-reports may be perfectly trustworthy, but *on average* they are likely to be roughly correct. On the

other hand, if one psychologist interviews a hundred men and if she happens to believe that there is a midlife crisis, she may well accept the stories of the distressed and dismiss the others as defensive distortion. Of course, if she has other biases, she may reach other conclusions equally inconsistent with the facts. The point is that interviewers are not necessarily impartial observers. Clinicians are trained to see pathology, and see it they will, even in the best-adjusted interviewees. Since theorists have a vested interest in finding evidence for the ideas they have advanced, the "data" they find through interviews in support of their own ideas must be taken with a certain degree of skepticism.

There are some precautions that can be taken to avoid the introduction of biases. For example, the interviews may be tape-recorded or videotaped and two or more judges can then rate the subject on the variables of interest. Agreement between them would be evidence that there is some basis in the interview for the inferences that were made and would seem to be a minimum requirement for taking the study seriously. Even perfect agreement, however, would not necessarily mean much. For example, a subject who was extremely nervous during the session might be characterized by both raters as anxious, although his behavior in the interview was unusual, the result of some temporary and extraordinary circumstance.

Further, an interview is an interpersonal encounter. The interviewer is not an uninvolved, dispassionate observer, but rather an active participant. As any courtroom lawyer will tell you, people can be led to say a great many things under the skillful questioning of an expert. In our courts of law, we try to compensate for this suggestibility by allowing a cross-examination in which a second lawyer, representing an opposite point of view, sees how far the testimony can be shifted in the other direction. In psychological research, this method is never employed (although it might be worth a try). When an adult developmentalist is looking for signs of a midlife crisis in a person, he must act as his own opposing counsel, cross-examining the subject to test all possibilities. Lawyers are not allowed to do this, since they are clearly motivated to help their clients. Researchers, on the other hand, are presumed to be motivated only by a desire for truth.

The pros and cons of the interview method reflect the larger issue that has always profoundly divided personality psychologists: the role of the unconscious. There are those who believe in the unconscious as the real heart of personality, the true source of action, meaning, and

value; and there are those who relegate it to a minor role or dismiss it entirely. Communication between these two schools is like an exchange between a believer and an atheist: The fundamental premises are so different that neither can understand the other's point of view. Like religion, the unconscious must in part at least be taken on faith. Once its existence and importance are accepted, however, it seems to make sense of many phenomena.

Researchers who prefer interviews are generally believers in the unconscious; they prefer the interpretations of an outside observer because they are fundamentally skeptical of the utility of self-reports. They assume that people are not only unaware of what is really going on in the important parts of their being but they are also highly motivated to remain unaware, since conflicts, anxieties, and unwelcome impulses predominate in the unconscious. If survey after survey shows no marked increase in depression or anxiety at the midlife (Lacy & Hendricks, 1980; Tamir, 1982), this is taken as clear evidence that surveys don't measure the real state of mind of individuals.

Researchers who rely on consciousness as the focus of personality are equally skeptical about the claims of interpreters. They point out that hypotheses about the unconscious are often irrefutable: Nothing we can do would disprove a dynamic interpretation. If, after several hours of intense probing, an interviewer is unable to find any conflicts that might give evidence of a midlife crisis, he or she can always conclude that the turmoil and despair are repressed—in fact, this must be an extremely severe crisis if it is necessary to repress it so thoroughly! Dire predictions of a future crisis are likely to be made, and if anything ever goes wrong in the person's life (as eventually something must), the prediction is trumpeted as proof of the theory.

By now, of course, it must be obvious to readers that our own biases favor conscious over unconscious processes and self-reports over outside interpretations of personality. Yet we are not prepared to dismiss entirely the contributions of a number of distinguished thinkers who have written extensively on the development of personality in adulthood. In reviewing their positions, we will certainly point out all the reasons we have for disputing many of their conclusions. But we will also try to point out the areas of agreement and the large areas in which their theories and research complement and enrich our own. As we will see, the substance of their findings can often be retained when it is modified or reinterpreted to make it consistent with the facts as we

know them. The result, we hope, will be a contribution to the study of emerging lives as well as of enduring dispositions.

Form and Content in Psychological Interviews

As it happens, most of the major theories of adult development are based on interviews rather than on psychological test results. The formulations of Erik Erikson and Roger Gould originated in their clinical practices; the theories of Daniel Levinson and his colleagues are based on biographies constructed by subject and interviewer. If these researchers have indeed discovered developmental courses standard measures cannot detect, the interview approach may be credited.

In our discussion of the interview method, we have so far been concerned only with the quality of the characterizations it produces. Is Jones an introvert? Do we believe it when he says he is? Or do we observe him during the interview and reach our own decision?

A review of theories of adult development shows, however, that theorists who rely on interviews do not have much to say about Extraversion. They are concerned with concepts such as the self and the sense of identity; they talk about relationships, roles, and life structure. This difference in substance between interviews and questionnaires is not coincidental. It seems to be a clear instance of form dictating content.

The typical questionnaire does not let you talk back. The questions must be phrased so that they can be meaningfully answered by all respondents with a choice from *yes* or *no*, *true* or *false*, or other fixed categories. If we are interested in job satisfaction, our questionnaire can ask, "Are you generally satisfied with your job?"; regardless of the kind of job one has, everyone who is employed can answer that question with a simple *yes* or *no*. It cannot ask, "What specifically do you like about your job, and what do you dislike?" since the answers to these questions would differ from one job to the next, and standard, precategorized answers would not suffice.

Again, if a researcher who uses questionnaires wants to know how an individual gets along with others, that researcher is likely to write questions such as "Do you have many friends?" or "Are you usually dominant in relations to others?" or "Do you believe others are out to get you?" By contrast, an interviewer is likely to say something like

“Tell me about your work. How do you get along with your bosses, co-workers, subordinates?” Or, “Describe your relationship with your wife and children.”

In some respects, we would expect the two methods to yield comparable results. After all, if an individual is really dominant, it should show up in the way he or she deals with friends, coworkers, and family. On the other hand, there are also respects in which important differences may emerge. Questionnaire results are standardized and uniform, independent of the content of the individual's life. Interview results are permeated by the details of an individual's life: the politics at the office, the specific problems of raising a teenage daughter or of living in a declining neighborhood. Case studies with all these details seem to provide a much better feel for the individual's personality than do raw test scores. But what they in fact give is *a better feel for the individual's life structure*, which may or may not be an accurate guide to personality. One's life structure is determined by countless forces beyond one's control: economic realities, accidents and illnesses, the help or interference of relatives. These are important elements in understanding the individual's world, but they are a source of potential confusion when one is trying to understand the individual.

INTERVIEW-BASED THEORIES OF ADULT DEVELOPMENT

At least one of the major theorists we will discuss, Daniel Levinson, acknowledges this important difference. He says explicitly that by adult development he means the “evolution of the life structure”—although he believes that there is an inner, psychological side to that process that seems akin to personality development. We believe this distinction is crucial to an understanding of the field: Nothing we have said about the stability of personality should be interpreted to mean that the life structure does not evolve, and the theories of adult developmentalists need not contradict our position on enduring dispositions. Agreement in principle, however, is not necessarily agreement in practice, and it will be necessary to look at each case closely to see what portion we are willing to believe. We return to the relation between personality traits and the life structure in Chapter 10 when we offer a theoretical framework for the Five-Factor Model.

Levinson's Seasons

The Seasons of a Man's Life (Levinson, Darrow, Klein, Levinson, & McKee, 1978), one of the most influential works on adult development, had its major impact before it was published. While he was still working on the book, Daniel Levinson discussed his ideas with Gail Sheehy, who went on to write the best-selling *Passages* (1976). Levinson and colleagues' research was based on 10–20 hours of interviews with each of 40 men in the age range from 35 to 45. Four occupational groups—executives, workers, biologists, and novelists—were selected to provide some variation, since occupation has a decisive effect on shaping the life course. The subjects were “normals”; that is, they were not recruited from hospitals or psychiatric clinics, although several had been in psychotherapy at some time in the past. (Much later D. J. Levinson and J. D. Levinson, 1996, published a work claiming to replicate their findings in women.)

In the course of their interviews with men, Levinson and his co-workers began to see a pattern. Instead of finding that the life course was shaped by external events like marriage, job promotions, or illness, they believed they had found evidence of a universal, age-linked series of stages, organized into broader eras. Childhood, Early Adulthood, Middle Adulthood, and Late Adulthood (and possibly Late-Late Adulthood) were the major eras, each lasting about 20 years, and each broken further into alternating stages of transition and stabilization. The major transitions, including the Midlife Transition at age 40–45, overlapped both the preceding and following eras.

The names given to the different stages summarize their developmental meaning. The Early Adult Transition (age 17–22) involves breaking away from home and facing adulthood. This stage had already been heavily researched (e.g., Constantinople, 1969), since for many individuals it occurs at college, where most subjects for psychological studies are recruited. Entering the Adult World comes next, as the man begins to function as an adult, usually starting a career and family. These initial attempts at acting like a grown-up are likely to be clumsy, because they begin without the benefit of experience and often with a premature haste to settle down and prove one's adulthood; so between ages 28 and 33 the Age Thirty Transition occurs, in which the early occupational and marital choices are reexamined and either rejected, modified, or reaffirmed.

Having made these adjustments, the man enters the period of Settling Down, in which he begins to take life seriously and sets foot on the ladder of success towards his Dream, the ambition he harbors deep down. He is likely in this period to come under the protection, guidance, or sponsorship of a mentor, a somewhat older man whom he admires and who encourages him in his Dream. His wife, also, must do her part to encourage him if he is to be successful. The mentor, who acts somewhat like an older brother, helps the man go from junior adult to senior, but in the latter years of the 30s Becoming One's Own Man becomes crucial, and achievement of occupational aspirations is often accompanied by alienation of the mentor.

Just when life seems to have peaked (at least for those who made it up the ladder), the Midlife Transition is entered and everything has to be reevaluated. If men have been successful in striving for their goals, they question whether the goals were meaningful. If they have not been successful, they must reconcile themselves to their failures. In addition, Levinson says, a number of other issues that emerge here must also be dealt with. Men must face the fact that they are getting older, a realization brought about by the beginnings of physical decline and perhaps by the death or disability of their parents. They must also resolve conflicts about masculinity-femininity, creation-destruction, and attachment-separation. In some men these conflicts are expressed as conscious introspection and philosophizing or in artistic creations; in most the conflicts are recognizable only as a sense of turmoil, stagnation, alienation, and confusion.

After age 45 the individual enters the Midlife period with a new basis for living and interpreting life. In the best of cases, concern over personal advancement and proving one's worth has been replaced by a more altruistic concern for the welfare of others and posterity (Erikson's *generativity*); and the single-mindedness of the 30s is replaced by a new perspective in which previously neglected aspects of the self are given expression.

Although he had not yet made a systematic study of men over age 50, Levinson assumed that similar stages will continue throughout life. The process of *life review* (Butler, 1963), once thought to be a characteristic of old age, may in fact simply be the last in a series of life reviews, undertaken at each transition point throughout life.

Levinson's theory, as he promised, is directly concerned with the life structure and only secondarily with personality. Yet there are two

points with important implications for the student of personality: First, to the extent that personality can be identified with values, concerns, and interests, it appears that major changes in personality are postulated to be the result of aging. Second, the turbulent nature of the life course as described by Levinson suggests that something like cyclical change in Neuroticism is involved. If individuals go through 5-year periods of anguish, conflict, and alienation every few years, measures of psychological distress ought to show it.

This theory is certainly more interesting than the bland assertion that personality is stable in adulthood; the question is, is it true? A careful and critical review of available evidence casts more than a little doubt on the entire scheme. These are among the problems:

Is It Universal?

Levinson claimed that the sequences he found are universal, at least in his sample. But reading the cases even as he presents them leaves doubts. Some people had ostensibly smooth transition periods; some tried to break out of their life structure when they should have been settling down (of course, they paid the penalty later for that transgression). Among the workers in particular, the success-ladder model did not seem to make much sense: In the United States there is generally not much hope for advancement for a 35-year-old factory worker.

Even if the scheme were descriptive for all Americans, the claim for universality seems premature. In support, Levinson offered the stage theories of Solon, Confucius, and the ancient Hebrews, which he believed showed parallels to his theory. But none of these three incorporates periods of crisis—surely the cardinal feature of Levinson's scheme—and one is on a 10-year cycle, another on a 7-year cycle. Confucius, for example, says that at age 40 he “no longer suffered from perplexities.” Does that sound like the beginning of a midlife transition?

What these timetables seem to indicate is a universal recognition that there are growth and activity in the first half of life, and reflection and decline in the second. They also illustrate the universal appeal of dividing up the seamless web of adult life into convenient, although essentially arbitrary, stages. The remarkable variety of “natural” divisions of the lifecycle is clearly shown in an essay on William Shakespeare's “seven ages of man” (S. C. Chew, 1947).

Is It Intrinsically Age-Related?

One of the most controversial features of Levinson's theory is his use of chronological age as the basis for his stages. Most other adult developmentalists prefer to break up the life cycle with marker events such as marriage, birth of the first child, and retirement. Sociologically, they argue, the newlywed 18-year-old male is more like the newlywed 28-year-old than like his bachelor age peers. Levinson counters by noting that the marriage of teenagers is likely to be as immature as they are; the event is colored by the period in which it occurs, not vice versa. On the other hand, the teenage father in this culture may be immature because he has not had the work and family experience of the 28-year-old (perhaps maturity comes sooner in less-developed cultures). Experience, not age per se, may be the crucial variable.

Is It Discontinuous?

The notion of stages implies qualitative differences between one period and another. Even if the changes that Levinson suggests actually come about, is there evidence that they do so in quantum leaps? Take, for example, the realization that one is growing older. Levinson is probably quite correct in pointing out that the fact of aging requires readjustment. Time goes by continuously, but our conception of ourselves does not change at the same rate. We tend to think we are as we have always been until something confronts us with our age and forces us to acknowledge the change. But it seems unlikely that this occurs in discrete, age-related stages. A thousand incidents may contribute to this insight: a first gray hair, a 20th anniversary or 40th birthday, the death of a friend or parent, a newsreel from one's childhood with grotesquely dated fashions and expressions, the observation that college students get younger and more immature each year. Each of these small realizations may be something of a jolt, but is there any evidence that they all occur together at precisely timed intervals? A man who still thought of himself as a teenager at age 39 probably would be in line for a midlife crisis, but are there many men like that?

Is It Personality Change?

When Levinson is confronted with evidence of stability in basic personality traits (Rubin, 1981), he tends to dismiss it. Dispositions such as

anxiety, gregariousness, or aesthetic sensitivity are trivial in his view, superficial aspects of temperament that do not address the real, deep level of personality. For him, the fact that a 50-year-old man views himself as middle aged when once he viewed himself as young is a startling change in personality; *young* and *old* are not merely age categories, and calling oneself middle aged does not mean simply that one is concerned about retirement plans and grandchildren. Young and old are symbols, Jungian archetypes carrying profound connotations of vitality, immortality, and springtime in contrast with death, decay, and winter. Acknowledgment of middle age is thus a change in the deepest symbolic significance of the self.

Certainly, with age and experience come changes in some aspects of how we view ourselves, and the idea of aging is obviously related at some level to the metaphor of life and death. Perhaps poets are moved by these metaphors to change the style of their poetry as they age. But does the change in the self-concept bring about important changes in the behavior or experience of most individuals? Does it lead them to depression and suicide or to altruistic self-sacrifice? The answer, we think, is "No." When children leave home, when aged parents require care, or when one retires from a lifelong occupation, there are profound changes in the daily routines that constitute the bulk of behavior, but these changes do not amount to changes in personality, and they come about in response to external necessity rather than internal development. Men do not give up playing tennis because they start to feel old; they are more likely to feel old because they have been forced to give up tennis. Nor is there much evidence, as we will soon see, that people feel different during or after one of the hypothesized transition periods.

Is It Developmental?

The field of gerontology has always been fundamentally ambiguous about the *value* of the changes it studies. University programs in this area are typically called "Aging and Human Development," as if the negative connotations of the first word had to be offset by more positive ones. Levinson usually writes as if what he is describing were development, a positive change, growth, and the attainment of higher unity and deeper wisdom. As he progresses through life, the young man learns more about himself and his world and, in the best of cases, is able to rise above the conflicts and illusions of youth to a greater maturity. He can

accept the feminine side of his nature and can acknowledge the individuality of his wife; his judgment is improved, marked by realistic compassion rather than by idealism or egocentrism. The older man sees the universal rather than the particular and is no longer bound by short-sighted or parochial views. With the abatement of instinctual drives, the older man can grow in rationality and ego strength. (Of course, Levinson would add, not every man succeeds in development: Some despair; some die too young. And the goal is never finally reached: Late life, like middle age, is marked by conflict, reassessment, and growth.)

On the other hand, in a discussion of this topic near the end of his book, Levinson expressly denies that his scheme is hierarchical, that later stages are better than earlier. Each has its strengths and weaknesses. But whereas no one stage is intrinsically better than another, the traversing of stages is itself the mark of growth. Individuals who do not change—who experience no periods of crisis—are arrested in development; they are, as a leading theorist of adult development, Orville Brim, said, “stuck” (Rubin, 1981). This view, however appealing and humanistic it seems, has as yet no basis in evidence. Change per se is not necessarily good, nor is stability necessarily the same as stagnation.

Gould's Transformations

If Levinson is fundamentally a Jungian theorist, Roger Gould is clearly a Freudian. He sees adult development as the dismantling of illusions of safety developed in childhood and maintained in a quasi-unconscious fashion through the first half of adulthood. His ideas were also borrowed by Gail Sheehy, and he too wrote a book, *Transformations* (1978), expounding his more fully formed theory.

Externally, there are both similarities and differences between Gould's and Levinson's work. Gould divides adulthood into five periods: from 16 to 22, when the adolescent is breaking ties with parents; from 22 to 28, when a new, adult life is started; from 28 to 34, when life is reassessed in view of conflicting internal needs and values; from 35 to 45, when the problems of evil, death, and destruction are faced; and postmidlife, when the individual is finally in control of his or her own destiny. Whereas Levinson calls for alternating periods of stability and transition, Gould sees a continually deepening—and progressively liberating—struggle. Gould expects no further development after 50, whereas Levinson does (note that neither of them studied individuals in

this age period). The time frames show some overlap, especially if Levinson's stable periods are ignored, but whereas Levinson views the timetable as inherent in human nature, Gould believes it is a function of life events and social forces (marriage, birth of children, corporate schedules of advancement, etc.). Gould also acknowledges the profound differences of social class. Perhaps the most distinctive feature of Gould's book was his effort to report on women as well as men and his particular attention to couples.

But the more fundamental differences are to be found in the theoretical underpinnings of *Transformations*. Gould sees the basis for adult development in the unfinished business of childhood. As children we are fundamentally helpless in the face of both outer danger and inner passions of lust, rage, and greed. We depend wholly on our parents to control both of these threats, and we internalize a series of false assumptions, illusions that allow us to believe that we are perfectly safe. Maintaining these beliefs has the benefit of preserving our sense of security, but it also has a cost: We are confined by the rules that bound us as children. We cannot get free of these confining inhibitions without facing the illusory nature of some of our most fundamental beliefs and without giving up the security they provide. But when we do get free, we come to see reality more clearly and thus stop the rude and unexpected shocks that must repeatedly occur when our illusions collide with life. And, says Gould, we also gain from this process real freedom to be our own persons, in touch with our inner needs and passions, living vital and meaningful lives.

Now, these ideas are by no means unprecedented. Most versions of depth psychology claim that neurotic hang-ups stem from defenses we adopted in childhood to deal with a reality that no longer exists. Even cognitively based theories of psychotherapy, like Ellis's (1962) Rational-Emotive Therapy, recognize the entrapping and self-perpetuating role of irrational beliefs. Gould's unique contribution is to insist that there is an *age-ordering* of these irrational beliefs. Teenagers entering adulthood confront the assumption that they will always belong to their parents and believe in their world. Achieving independence from one's parents and their values, taking charge of one's life and one's body, means giving up the comforting assurance that parents will always be there to help and guide.

The young adult who is building a family and beginning a career has joined the establishment—and, says Gould, has done so with a ven-

geance. The roles of adult, parent, husband, woman, and breadwinner are adopted wholesale, in part because the inexperienced adult has no other basis for guidance and in part because he or she believes that following all the rules, working hard, and persevering will guarantee happiness. To be sure, these kinds of activity may well be the best bet for creating a satisfying life, but no behavior can *guarantee* happiness, love, or success, and that is the lesson which must be learned, on an emotional as well as an intellectual level.

By age 28 a new phase starts. The difficulties of building a life tend to have been mastered well enough, and the internal side now clamors for attention. Wishes, values, and feelings that were ignored during the early pragmatic years of getting started suddenly take on a new importance. Career and marriage choices are reevaluated, and we come to the painful conclusion that life is not the simple, controllable affair we imagined. The illusions that we knew ourselves and that we were rational, consistent, and independent people are challenged.

Finally, in the midlife decade, an even harder pill must be swallowed: We must admit to ourselves that evil, death, and destruction are a real part of the world and even a real part of us. We have an evil side that may be controlled but cannot be eliminated; and we face a certain death, with time the only question. The decline of our parents at this time plays a part in shaking the belief that they can save us from ourselves or from external danger. Facing these facts may lead to a period of intense struggle and distress, a midlife crisis. But once the challenge has been confronted, once we have come face to face with our existential aloneness, we achieve the freedom that comes from what Gould called "owning ourselves."

How do we evaluate this theory? As a theory of psychotherapy, its first test would be in its effectiveness in helping clients to resolve their problems. In the hands of a skilled clinician such as Gould, the theory seems to be effective—but the history of psychotherapy teaches us that, in the hands of a skilled clinician, almost any theory can be effective. More convincingly, research Gould (1972) conducted on a large sample of adults who were not patients confirmed that there is a change in the "march of concerns" over the early adult lifespan: Young adults are concerned with getting along with (or away from) their parents; men and women in their 20s are concerned with starting family and career; those in their 30s complain more of stagnation; and for those at midlife, facing death and decline are relatively more frequent problems.

This is important information, although not really surprising. And a shift in the focus of concerns is something quite different from the evolution of adult consciousness. We might argue, for example, that irrational beliefs would be a dominant feature of the consciousness of only a small minority of individuals—those who require psychiatric treatment. We might grant that all individuals show some of the insights outlined by Gould, but we maintain that these changes are trivial readjustments in thinking and not the stuff on which life decisions are made. We might contend that the order of the stages is essentially arbitrary, a result of the most frequent course of events encountered by Americans. A 20-year-old diagnosed as having cancer may deal with issues of destruction and death before taking on other illusions. If this is the case, then *development* seems to be an inappropriate word. Instead, we might view the whole process as one of adaptation to the challenges of living as an adult.

Because Gould's theory is more elastic than Levinson's, it is harder to test. But one of the deductions that can be made from it is that there ought to be notable differences between the young and the old in basic adjustment and also in openness to feelings. As Chapter 4 showed, there is mixed support for that hypothesis: Adults are indeed lower in Neuroticism than adolescents, but they are also lower—not higher—in Openness. A second deduction would be that there are specific periods in adulthood when the struggle with irrational beliefs is particularly strong and overt crisis is likely. For Gould, this is most likely to be in the midlife period, from 35 to 45. We consider that possibility next.

IN SEARCH OF THE MIDLIFE CRISIS

Of all the features of adult development, the most celebrated is the midlife crisis. Jacques (1965) studied the lives of artists and decided that they all went through a period of crisis precipitated by the recognition of their own mortality. "Time since birth" was replaced by "time left to live" in the mind of the middle-aged man. Peter Chew (1976) wrote a popular account of the crisis, suggesting that marriage stales and the quest for a lost youth leads men to intensified sexual yearnings. Among women, menopause and the departure of children from home were viewed as likely causes of a period of depression and crisis. Levinson et al. (1978) wrote that "for the great majority of men . . . this period

evokes tumultuous struggles within the self and the external world" (p. 199) that may appear like neurosis to an outside observer.

These ideas have become so prevalent in both the popular and the scientific press that most people take it for granted that around age 40 there is a marked increase in such events as divorce and separation, suicide, job change, and admission to psychiatric hospitals. But epidemiologists who have scrutinized the figures on this issue find very little support for a period of crisis (M. Kramer & Rednick, 1976). Divorce is most common in the 20s; suicide is most common among the young and the old. Admissions to psychiatric hospitals show no peak at age 40. Still, these are all very rough markers of crisis; perhaps the midlife transition leads to more subtle manifestations.

When Levinson, Gould, and others were first presenting their views on adult development in the early and middle 1970s, the case for stability in personality had not yet been made. The midlife crisis had been pointed out so often that we, like most other psychologists, assumed that it occurred. The major question, we thought, was to determine its exact timing. Was it between 37 and 40 or between 40 and 45? Was the time different for working-class and middle-class men? We used our usual questionnaire methods to try to find an answer. Drawing on the literature describing the characteristics and concerns of men at midlife and the items that Gould (1972) had reported to differentiate patients of different ages, with Madeline Cooper we created a Midlife Crisis Scale and administered it to about 350 men aged 30 to 60 (Cooper, 1977). The scale had items covering sense of meaninglessness, dissatisfaction with job and family, inner turmoil and confusion, and sense of impending physical decline and death.

When we contrasted midlife groups with pre- and postmidlife groups, using various age cutoffs to define the midlife, we were surprised to find no evidence at all of a peaking of midlife concerns at any age in our range. We had also asked our men to describe in their own words how their lives were going just then. Only seven men (about 2%) seemed to fit the category of crisis, and these ranged in age from 34 to 56, distributed randomly over the age range we studied. Our conclusion: At any given time, only a small percentage of men are in a crisis, and they are not likely to cluster at any particular age.

Since this conclusion flew in the face of the rest of the literature, our first concern was to replicate it. With a different group of about 300 men, we used a shortened version of the Midlife Crisis Scale. The re-

sults were an exact reconfirmation: There was not the slightest evidence of a peaking of distress or midlife characteristics anywhere in the age range we studied (Costa & McCrae, 1978).

Again, the most persuasive data may come from the NHANES survey (Costa, McCrae, et al., 1986). Short scales measuring Neuroticism, Extraversion, and Openness were administered to nearly 10,000 men and women; about half of the respondents were in the age range from 34 to 54. With so many subjects, we were able to plot personality scores year by year. A crisis or temporary alteration in personality should be apparent as a spike in the graph, yet Figure 8 shows that neither men nor women showed any pronounced peaking at any point during the period. It is particularly noteworthy that there is no increase in Neuroticism around age 40, when the theories call for a period of emotional upset and psychological crisis.

Our disenchantment with stages of adult development can be traced to such results, but we are not alone in our reappraisal of the crisis. Farrell and Rosenberg (1981), who had once been proponents of the theory, undertook a major study of the midlife period. They commissioned a probability sample of about 500 men in the age ranges 25–30 and 38–48, a sample from which generalizations about American men could be made with confidence. They put together a battery of measures of alienation, distress, depression, and crisis, including a midlife crisis scale much like ours in content. They also had measures of what they called *authoritarian denial*, which is likely related to closedness to experience.

Their analyses, dividing subjects by age and by social class, showed almost exactly what we had: Midlife men were slightly higher in authoritarian denial, but they were actually a bit lower in alienation. There was no difference at all on their midlife crisis scale.

But perhaps all these studies make the mistake of tying the crisis to age when stage of life is more important. Among 40-year-old men, some are only recently fathers, many have young children in the home, many others have teenagers, and some have seen all their children grow up and leave home. Perhaps the crisis occurs during one of these stages instead of at a particular chronological age. Indeed, there is evidence from large-scale epidemiological studies (Tamir, 1982) that men with teenage children are a bit less happy than other men (hardly a surprising finding for those acquainted with teenagers). But most theorists had supposed that the real crisis would come when the children left home, leaving

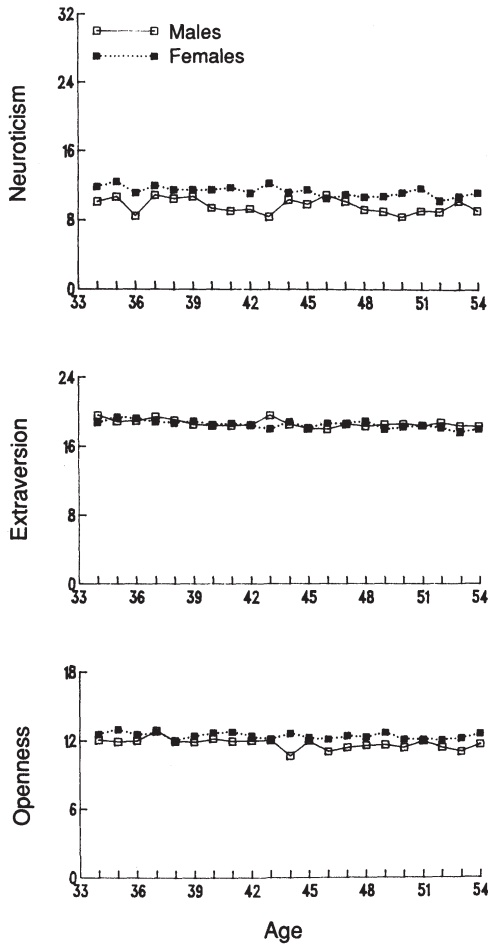


FIGURE 8. Mean levels of Neuroticism, Extraversion, and Openness to Experience for 1-year age groups of men and women aged 35–54. Adapted from Costa, McCrae, et al. (1986).

parents to face each other again, the stark silence reminding them that they were growing old. Some parents do react badly to the departure of their sons and daughters, but in one of the few systematic studies of this stage in the family Lowenthal and Chiriboga (1972) reported that, if anything, parents were somewhat happier in the empty nest.

“But this proves nothing!” say the critics. “These studies all employed paper-and-pencil tests. What is needed is the insight, probing,

and sensitivity of a trained interviewer. She or he would cut through the facade of well-being and see the crisis brewing just beneath the surface, ready to explode at any time into depression, suicide, divorce, alcoholism, dramatic career shifts."

Perhaps. Levinson's interviewers seem to have seen these things. But Farrell and Rosenberg (1981) also conducted interviews with 20 of the men they had surveyed, and they came to quite different conclusions. Many men had difficulties in life, but there was no evidence that problems were concentrated at certain periods. Further, the life histories these interviewers compiled from several hours of talking with the men and their wives and children convinced them that, for the minority who appeared to have a crisis, "the difficulties they experience have their roots in conflicts and problems of earlier origin" and are likely to lead to a "more general process of depressive decline" (p. 215). By contrast, they said of one man who showed Openness and positive adaptation in the middle years that "the sense of self he experiences at midlife is not a marked change from that of early adulthood." In search of a midlife crisis, these researchers emerged from intensive interviews with an impression of stability in the life course.

* * *

The interview, it seems, provides more ambiguous accounts of the period of adulthood than do standard personality tests. Some interviewers come to the conclusion that development proceeds in stages; some see no stages at all. There is even greater disparity when the stability of objective test results is contrasted with the growth and development described by interviewers. In part, this disparity results from the flexibility of the interview as a method of gathering data: If it is more sensitive to the subtle thoughts and feelings of the subject, it is also more sensitive to the biases and preconceptions of the interviewer. In part, however, it also results from the fact that interviews are typically filled with concrete details from the individual's life, not generalizations about dispositions. Lives surely change, perhaps in stages; personality traits, by and large, do not.

We see little reason to adopt the model of development proposed by Levinson and his colleagues, or the more widely held theory of a midlife crisis, and we are skeptical that the processes described by

Gould are universal developmental changes in personality or consciousness. But we must admit that the simple statement that personality is stable does not begin to do justice to the full complexity of adult life. There is both stability and change in life, and a conceptual framework that can accommodate both is needed. The time has come to provide a more formal account of the relations between personality traits, the self, and the life structure. We call it Five-Factor Theory.

A Five-Factor Theory of Personality

In common speech, a *theory* is a guess with little or no supporting evidence—as when a detective has a “theory” that the jilted lover is the murderer. When proof is found, the theory becomes a fact. In science, *theory* has a very different status. A scientific theory is a broad and abstract account of the general principles that are thought to explain a set of phenomena, like the General Theory of Relativity or the Theory of Evolution. Theories attempt to explain what is known in as simple a way as possible; ideally, they also make predictions about what is currently unknown. When those predictions are verified by observation or experiment, the credibility of the theory is strengthened.

In the natural sciences, theories are often stated as mathematical laws and quite specific predictions can be made. In the social sciences such precision is rarely seen, and in personality psychology, in particular, theories serve a rather different function. A theory of personality is a way of accounting for what people are like and how they act; a good theory explains a wide range of observations and points researchers in the right direction for future research. Freudian theory pointed researchers toward the study of dreams, but decades of research have yielded very little by way of supportive evidence (Domhoff, 1999). Trait theory pointed researchers toward general styles of thinking, feeling, and acting, and has resulted in thousands of interesting and useful find-

ings. That is why most personality psychologists today prefer trait theory to psychoanalysis.

Because theories give the “big picture” and guide thinking toward important issues, books like this one often begin with a theoretical statement as a background to the data to be discussed. And in fact we did introduce trait theory in Chapter 2; without knowing what traits are, it would make little sense to read about their stability over time. But trait theory as we have described it is not sufficient as a theory of personality, for two reasons: It says too little about traits, and it says far too little about nontrait aspects of personality.

Postulating that there are enduring and consistent individual differences does not go very far in accounting for traits. What are these traits? Where do they come from? How do they influence behavior? How can they be assessed? There are, of course, many answers to these questions, each corresponding to a particular version of trait theory. Indeed, one could say that what we have called trait theory is in fact a whole class of theories, some closer to the truth than others.

But even if we had answers to all those questions, we would not have a complete theory of personality, because there is more to human personality than traits. For example, in Chapter 8 we discussed functions usually attributed to the ego, which have to do with resolving conflicts, creating schedules, and overseeing the integrated operation of all the needs and tendencies of the individual within the constraints of the social environment. Anyone who has read a survey of personality theories can think of many other topics that are not traits: attitudes, roles, identities, the self-concept, moral development, and so on. A reasonably complete personality theory has to say something about all these.

The Birth of a Theory

Many theories in personality resulted from an effort to explain clinical phenomena, such as cases of hysteria (Breuer & Freud, 1895/1955) or students' difficulties in adjustment (Rogers, 1951). These are often called *armchair theories*, because they were the product of informal observation and thought, rather than scientific measurement and experimentation. There is nothing wrong with starting with informal observation; in fact, it is essential in the early stages of a science when it is not yet clear what should be measured. But personality psychology has been

around for well over a century, and contemporary theories ought, we think, to begin with the established facts. It was not until we had learned a great deal about how traits function that we ventured our own theory (McCrae & Costa, 1996, 1999).

The first and central fact that we had to deal with was the stability of personality traits over long periods of time. Again and again, our assertion that personality was stable was met with incredulity: How could it possibly be true that personality remains the same, when life presents a continuing stream of new experiences, new relationships, new opportunities and limitations on what we can do? It gradually occurred to us that we used the word *personality* differently from many others, and that traits (our principal interest) had to be sharply distinguished from relationships, habits, self-images, and so on. Something can change *and* stay the same only if it has at least two distinct parts.

Oddly, that insight was the basis of not one but two new theories of personality presented at a day-long 1992 American Psychological Association symposium on the topic "Can Personality Change?" Dan McAdams (1992) argued that personality should be construed as occupying three levels: Dispositional Traits, Personal Concerns, and Life Narratives. He conceded that traits are largely stable but pointed out that personal concerns (current goals, plans, worries, etc.) and life narratives (people's self-concept expressed as a life story) change with the situation and with the point in the lifespan. McAdams regards these as essentially independent levels, each requiring its own methods of study.

We (Costa & McCrae, 1994) presented a somewhat more ambitious model that also distinguished between enduring and changing parts of personality. Eventually (McCrae & Costa, 1999) we came to regard these parts not as the separate levels of McAdams's model but as interconnected components of a *personality system*. (At the same time, Mayer, 1995, was developing a systems-theory-based personality model that shows some resemblance to ours; Mischel and Shoda, 1995, have also, and independently, offered a somewhat similar systems model.) We also articulated a series of postulates intended to describe how the components acted and interacted with each other. We called the whole construction a *Five-Factor Theory* (FFT) of personality, because it was based on the whole body of findings associated with research on the Five-Factor Model (FFM). FFT began with observations of trait stability, but it is supposed to explain much more.

FIVE-FACTOR THEORY

Components of the Personality System

Figure 9 represents the components of the personality system according to FFT. The three rectangles refer to central components; the three ellipses to peripheral components that mark the interface with systems outside personality per se. *Biological Bases* and *External Influences* are the inputs, representing interactions of personality with the physical body and with the environment. The *Objective Biography* is the output: It is everything the person does, thinks, or feels across the whole life-span. It differs, of course, from the life narrative, which is a subjective biography that may be inaccurate and must be highly selective.

The three central components are labeled *Basic Tendencies*, *Characteristic Adaptations*, and the *Self-Concept*. As Figure 9 indicates, the Self-Concept is in fact a Characteristic Adaptation, but one of such interest and importance to psychologists that it has been given the status of a separate component. The heart of the model is the distinction between Basic Tendencies and Characteristic Adaptations, precisely the distinction that we need to explain the stability of personality. Basic Tendencies are the abstract capacities and tendencies of the individual, whereas Characteristic Adaptations are concrete acquired structures that develop as the individual interacts with the environment. Thus Basic Tendencies can be stable, while Characteristic Adaptations change.

Although FFT focuses chiefly on personality traits, Basic Tendencies also include cognitive abilities, artistic talents, sexual orientation, and the whole psychological machinery underlying learning, perception, and other psychological functions. For example, the capacity to learn language is a Basic Tendency that all human infants possess. But knowledge of French or Swahili or Quechua is not a Basic Tendency, even in those born to the culture; it is a Characteristic Adaptation. All learned skills are Characteristic Adaptations, as are habits, interests, attitudes, beliefs, and the internalized, psychological aspects of roles and relationships. Some Characteristic Adaptations, like language, are extraordinarily long-lived; others, like the personal concerns McAdams assigns to his second level, may be much more transient.

Although the distinction between Basic Tendencies and Characteristic Adaptations is not customarily emphasized in personality psychology, it is not particularly controversial. The claim that personality traits

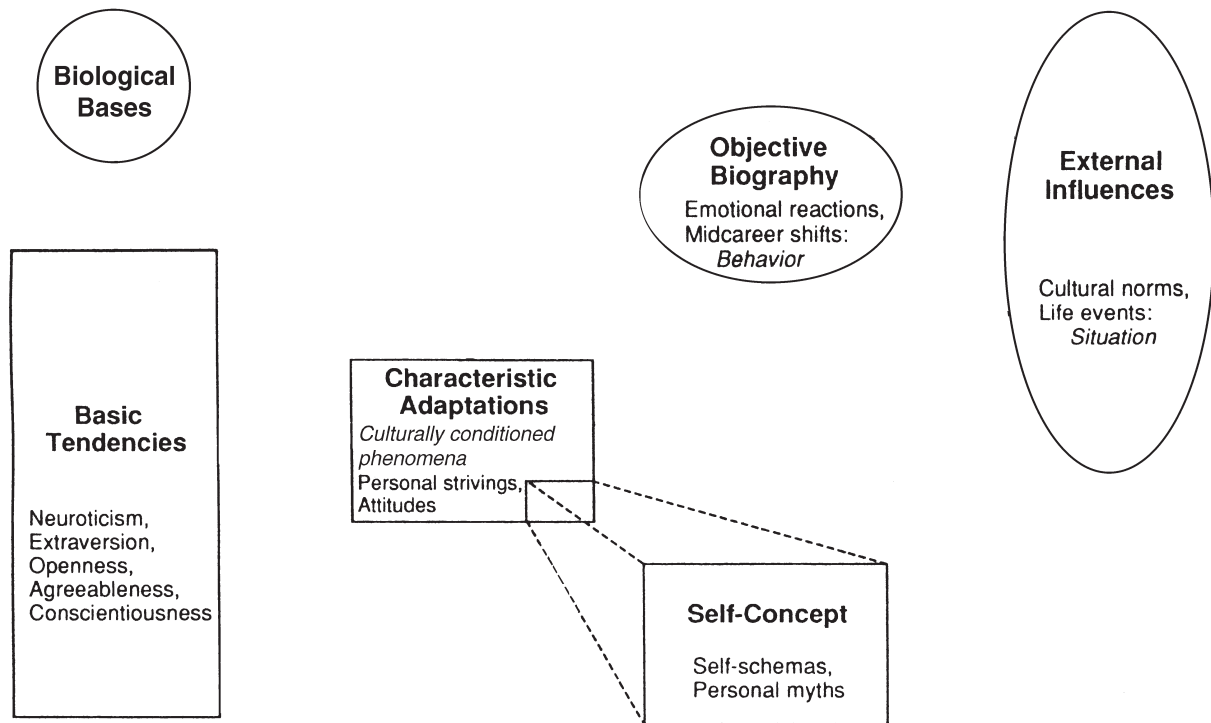


FIGURE 9. Components of the personality system according to FFT. Rectangles represent core components; ellipses represent peripheral components. Adapted from McCrae and Costa (1996).

are Basic Tendencies rather than Characteristic Adaptations, however, is. Many theorists have assumed or implied that traits are acquired, like broad and generalized habits. Cloninger, Przybeck, Svrakic, and Wetzel (1994), for example, distinguish between temperament and character. Although they might assign temperament traits to the category of Basic Tendencies, they would probably consider character traits to be Characteristic Adaptations. (Empirically, it seems to be hard to distinguish between temperament and character traits; both act like Basic Tendencies; see Herbst, Zonderman, McCrae, & Costa, 2000). We will return to the defense of this aspect of FFT later; for the present, we need to finish describing the theory.

As we will shortly see, there is another “component” to the FFT model, consisting of *Dynamic Processes* that regulate the interaction of the components in Figure 9. There are many such processes—for example, perception, coping, role playing, reasoning, long-range planning—but FFT says little about them. Something is happening; for the present, we are only interested in gaining an overview of the operation, not the details of how it works. In any case, many of those details will have to be supplied by specialists in such fields as cognition, neuropsychology, and attitude change.

Postulates of FFT: How the System Works

Table 16 lists the postulates of FFT as first formulated in 1996. We need not be concerned with all of these postulates, but some of them require comment. The Origin, Development, Adaptation, and Interaction postulates are most central to understanding how we think personality operates. In considering them it is useful to draw a picture, and Figure 10 serves that function. The arrows indicate the direction of influences, suggesting the operation of (unspecified) Dynamic Processes.

Figure 10 adds the dimension of time to Figure 9: It shows what leads to what. Figure 10 operates on different time scales—that is, it can be seen as a diagram of what is going on at any given moment, or it can be seen as a diagram of how the individual develops over time, even the whole lifespan.

Let us begin at the end, the output, the Objective Biography. This consists of the cumulative record of everything the individual does and experiences. What an individual does at any given moment, is, of course, *behavior*, the phenomenon many psychologists believe we are

TABLE 16. FFT Postulates

-
1. Basic Tendencies
 - 1a. **Individuality.** All adults can be characterized by their differential standing on a series of personality traits that influence patterns of thoughts, feelings, and behaviors.
 - 1b. **Origin.** Personality traits are endogenous basic tendencies.
 - 1c. **Development.** Traits develop through childhood and reach mature form in adulthood; thereafter they are stable in cognitively intact individuals.
 - 1d. **Structure.** Traits are organized hierarchically from narrow and specific to broad and general dispositions; Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness constitute the highest level of the hierarchy.
 2. Characteristic Adaptations
 - 2a. **Adaptation.** Over time, individuals react to their environments by evolving patterns of thoughts, feelings, and behaviors that are consistent with their personality traits and earlier adaptations.
 - 2b. **Maladjustment.** At any one time, adaptations may not be optimal with respect to cultural values or personal goals.
 - 2c. **Plasticity.** Characteristic adaptations change over time in response to biological maturation, changes in the environment, or deliberate interventions.
 3. Objective Biography
 - 3a. **Multiple determination.** Action and experience at any given moment are complex functions of all those characteristic adaptations that are evoked by the environment.
 - 3b. **Life course.** Individuals have plans, schedules, and goals that allow action to be organized over long time intervals in ways that are consistent with their personality traits.
 4. Self-Concept
 - 4a. **Self-schema.** Individuals maintain a cognitive-affective view of themselves that is accessible to consciousness.
 - 4b. **Selective perception.** Information is selectively represented in the self-concept in ways that (i) are consistent with personality traits and (ii) give a sense of coherence to the individual.
 5. External Influences
 - 5a. **Interaction.** The social and physical environment interacts with personality dispositions to shape characteristic adaptations and with characteristic adaptations to regulate the flow of behavior.
 - 5b. **Apperception.** Individuals attend to and construe the environment in ways that are consistent with their personality traits.
 - 5c. **Reciprocity.** Individuals selectively influence the environment to which they respond.
 6. Dynamic Processes
 - 6a. **Universal dynamics.** The ongoing functioning of the individual in creating adaptations and expressing them in thoughts, feelings, and behaviors is regulated in part by universal cognitive, affective, and volitional mechanisms.
 - 6b. **Differential dynamics.** Some dynamic processes are differentially affected by basic tendencies of the individual, including personality traits.
-

Note. Adapted from McCrae and Costa (1996).

supposed to explain. Kurt Lewin famously said that behavior is a function of the person and the environment: $B = f(P, E)$; but FFT has a slightly different answer. Figure 10 and the Interaction postulate claim that behaviors result from the interaction of Characteristic Adaptations and External Influences: $B = f(CA, EI)$. For example, Joan might be offered tickets to see *La Traviata* (an External Influence), but she has a long-standing aversion to grand opera (a Characteristic Adaptation). She therefore declines the offer (the Objective Biography). In practice, even the simplest behavior involves more than this brief account suggests. Joan must understand the language the offer is made in, she must consider her relationship to the individual making the offer, she must remember her feelings about opera (a part of the Self-Concept), she must follow the dictates of etiquette in declining the invitation, and so on. It is not possible to behave as a human being without having a vast and intricate network of skills, beliefs, and values that allow one to interact meaningfully with the social environment. That network is what FFT calls *Characteristic Adaptations*.

Figure 10 shows that the behavior itself has consequences, feeding back on the Self-Concept (strengthening Joan's sense of herself as a country-and-Western fan) and on the environment (that's the last invitation she'll get from that friend!). But nothing we have described so far has anything directly to do with traits. To involve them, we need to invoke the Interaction and Adaptation postulates. These postulates specify the origins of Characteristic Adaptations in the interaction of Basic Tendencies and External Influences: $CA = f(BT, EI)$.

In fact, this is why they are called Characteristic Adaptations: They are *characteristic* because they reflect the operation of enduring personality traits, and they are *adaptations* because they are shaped in response to the demands and opportunities offered by the environment. Personality traits have more influence on some Characteristic Adaptations than others. Whether one eats with chopsticks or silverware is almost entirely a function of where one was born. But even here, personality can intrude: More adventurous, inquisitive, open people are more like to learn foreign customs, like eating with silverware.

To continue our earlier example, we can now account for Joan's aversion to opera. That Characteristic Adaptation is probably the result of her life experience and her fundamental closedness to Aesthetics. If she is very closed, she may never even have heard an opera and has formed an opinion based solely on reputation. (The Baltimore Opera once used the advertising slogan, "It's not as bad as you think—it can't

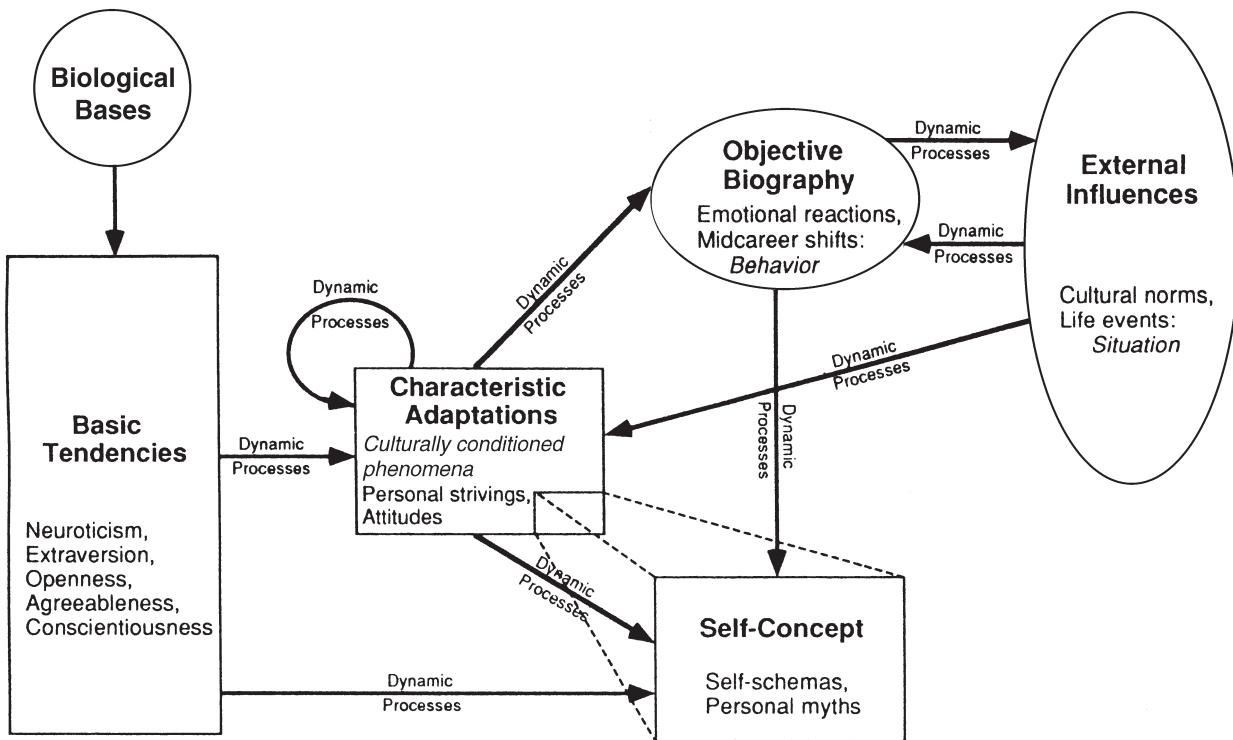


FIGURE 10. Operation of the personality system according to FFT. Arrows indicate the direction of causal influences, which operate through dynamic processes. Adapted from McCrae and Costa (1996).

be.”) Or perhaps she has been to an opera or two, found nothing of interest, and concluded that opera was not for her. Being closed, she has discovered that she is most comfortable being around down-to-earth, conventional people (McCrae, 1996), and they share and encourage her musical tastes. Her underlying closedness has influenced a host of Characteristic Adaptations, which generally reinforce each other.

The Origins of Traits

Where, then, do Basic Tendencies, specifically personality traits, originate? Here we come to the most controversial aspect of FFT. The Origin postulate says that traits are endogenous Basic Tendencies, and in Figure 10, the only arrow leading into the Basic Tendencies component is from Biological Bases. In contrast to virtually all other personality theories, FFT does not admit of any influence of the environment on personality traits (McCrae et al., 2000). This is a very radical position, and it is probably wrong in an absolute sense—that is, we are likely to find circumstances in which the environment does have direct influences on traits. Our intent, however, is to sketch a parsimonious model that accounts reasonably well for known facts, and we believe FFT does that.

The original basis for this position was, once again, the stability of personality. We know that life changes for everyone over time, and it changes dramatically for some people. If traits are shaped by the environment, we should expect considerable change in traits over time, as most adult developmentalists did 20 years ago. But we find very little change. Therefore, traits must *not* be shaped by the environment.

No one—including us—would accept that conclusion if the only evidence were from longitudinal studies of personality stability. A single source of evidence is rarely enough to support so broad an interpretation. But there are several other lines of evidence that make Figure 10 more plausible.

Genetics

Consider first data on heritability. There is by now a very large body of data on the genetics of personality traits, from adoption studies (Loehlin, 1992) to studies of twins raised in the same family (Jang et al., 1996), to studies of twins separated at birth (Bouchard & McGue,

1990). All of these studies are remarkably consistent. They suggest that about half the variance in personality traits scores is attributable to genes and that almost none is attributable to a shared family environment. The “shared environment” consists of all the influences that children share by growing up in the same family, including parental models, family values, child-rearing practices, local schools, diet, religion, and so on. According to the behavior genetics literature, none of that seems to matter in shaping the personality traits of adults. That conclusion is supported by studies of fraternal and identical twins, but it is seen perhaps most clearly in adoption studies. When biologically unrelated children are raised in the same family, they show no resemblance whatsoever in personality as adults. They may follow the same religion or speak with the same accent or prefer the same kinds of food, but they are no more alike in levels of Extraversion or Conscientiousness than two people plucked at random from the population.

The nonimportance of the shared environment had become clear to behavior geneticists by the mid-1980s (Plomin & Daniels, 1987), and many efforts have been made to understand it (Pike, Manke, Reiss, & Plomin, 2000), so far with limited success. If genes account for half the variance and the shared environment none of it, what accounts for the other half? Technically, the remainder is called the “nonshared environment,” and some researchers interpret it as all those shaping events that distinguish two children from the same family. For example, both Billy and Judy may attend the same local elementary school, but Billy is taught by Ms. Smith and Judy by Ms. Jones. Or perhaps Billy takes piano lessons while Judy studies guitar. Or Billy sits on the right side of the dinner table, whereas Judy sits on the left. In principle it is possible that such differences might have important influences on personality development, but it does not seem very plausible. Why should the choice of instrument make a difference when it does not seem to matter that Billy and Judy are both given a musical education?

Judith Rich Harris (1998) created something of a sensation when she suggested that family influences count for much less than peer influences. We know that many children of immigrants grow up hearing their ancestral language at home and English at school, and when they grow up, they speak English. Perhaps in the same way, children acquire personality traits from their peers rather than their parents. Because children from the same family generally have somewhat different peer groups, they may show limited personality resemblance as adults.

This is a plausible hypothesis, but there is as yet little evidence by which to evaluate it. If we could identify cliques of children who stayed together through most of childhood, we might hypothesize that, as adults, they would show more resemblance in personality traits than children from different cliques. We do know that people who grow up in the same culture tend to show similar personality profiles relative to people from other cultures (McCrae, 2000a), but it is not clear how to interpret that finding. After all, people from the same culture also tend to be genetically more similar than people from different cultures.

There is another interpretation of the nonshared environment, consistent with FFT. Perhaps it is mere error of measurement. We know that all personality assessments include some degree of random error, reflected in the reliability (or unreliability) of the test. In addition, however, personality assessments include systematic error that can lead to misinterpretations of data (McCrae, Jang, Livesley, Riemann, & Angleitner, 2001). Systematic error or bias cannot be distinguished from valid variance when only one source of information is used. Heinz may overestimate his Conscientiousness, but if he does so consistently, it may seem that his estimate is correct. It is only by asking his wife and his friends, or by observing him ourselves, that we can figure out whether and by how much he is biased. When Riemann and colleagues (1997) based their estimates of heritability on a combination of self-reports and peer ratings, they found substantially higher levels of heritability than are usually reported, ranging from .66 to .79 for the five factors.

Other Biological Influences

These unbiased estimates suggest that genetics, not the environment, is responsible for most of the variance in adult personality traits—but not all. That does not mean that FFT is wrong, though, because there is more to Biological Bases than genetics. Finch and Kirkwood (2000) point out that chance plays a large role in human biology. Genes provide a blueprint, but in the actual course of development—especially embryonic development—that blueprint is not followed precisely. Mutations, intrauterine hormones (Resnick, Gottesman, & McGue, 1993), and other factors subtly affect the finished product. That is why “identical” twins are never quite identical, even in such purely physical attributes as height or fingerprints. There is no reason to expect that personality traits would be different in this regard.

Several other biological mechanisms clearly affect personality. In the 19th century Phineas Gage survived an accident that left a metal shaft lodged in his brain. He was observed afterward to have had profound changes in personality traits, becoming irritable and impulsive. Neurologists have documented thousand of cases of traumatic brain injury since then, with attendant changes in personality (Costa & McCrae, 2000a). Similarly, disease processes can affect the brain and, through it, personality traits. Siegler and colleagues (1991) were the first to show the effects that dementing disorders like Alzheimer's disease have on personality traits—most notably, a precipitous decline in Conscientiousness. Mental disorders like schizophrenia and acute depression also affect personality traits, and pharmacological treatments can reverse many of these effects (P. D. Kramer, 1997).

The astute reader will note that some of these examples illustrate a sense in which the environment can be said to affect Basic Tendencies: by altering the underlying biology. Figure 10 might well be modified by adding another arrow between External Influences and Biological Bases. That would add another important pathway without affecting the basic distinctions drawn by FFT.

Conflicting Evidence

Although the Origin postulate of FFT appears to have considerable support, it must be pointed out that there are several bits of evidence that contradict it by suggesting a direct effect of the environment on personality traits. Clinical depression can be successfully treated by conventional psychotherapy as well as by drugs (VandenBos, 1986); presumably that therapy also affects personality traits. Life events per se usually do not change personality trait levels, but some specific events apparently do. For example, Costa, Herbst and colleagues (2000) found that women who were divorced in the 6-year period before a longitudinal readministration of the NEO-PI-R became somewhat more open and extraverted compared to women who got married in the same period, as if divorce were a liberating event—and marriage a constraining event. It was not clear, however, how long lasting the changes would be, because there was no follow-up. Again, an acculturation study of Hong Kong Chinese undergraduates who had moved to British Columbia suggested that residence in Canada led to higher levels of Openness and Agreeableness (McCrae, Yik, Trapnell, Bond, & Paulhus, 1998).

Brent Roberts (personal communication, August 1, 2001) believes the longitudinal literature contains a number of documented instances of life experiences that caused meaningful changes in personality traits. We are not yet persuaded that these effects are robust and replicable or that they really refer to changes in Basic Tendencies. Consider the acculturation study (McCrae, Yik, et al., 1998). Hong Kong Chinese scored lower than Americans on measures of Altruism, because China is a collectivistic society in which individuals' predominant loyalty is to their family and business associates. So much is demanded of Chinese by these in-group members that little is left for out-group members. In consequence, Chinese respondents score low on the NEO-PI-R Altruism scale, because it assesses altruism mainly by items like "I think of myself as a charitable person." Chinese who move to Canada may learn another set of values (a Characteristic Adaptation) in which charity to out-group members is deemed appropriate. Thus, acculturation may simply change the way in which Altruism is expressed.

We would not claim that we have convincingly demonstrated that External Influences have no effect on Basic Tendencies, but we think the burden of proof has shifted 180 degrees from the conventional wisdom of psychologists: It is now incumbent on those who think the environment shapes personality to prove it. More realistically, since science is rarely a matter of either one principle or its opposite—either biology *or* environment—what is needed is a specification of the particular conditions and circumstances in which life experience can reshape personality and an explanation of how it does so. The great majority of life events do not have lasting effects on personality traits; what is it about those few events that do that sets them apart?

Trait Development and the FFT

Although FFT does not give a role to life experience in shaping personality traits, it does acknowledge that trait levels change across the lifespan. As it was originally formulated in 1996, FFT claimed that *traits develop through childhood and reach mature form in adulthood; thereafter they are stable in cognitively intact individuals*. By 1999, we acknowledged a need to modify that statement, because evidence suggested very small but consistent changes in Neuroticism, Extraversion, and Openness after age 30. We now have some idea of how traits develop from adolescence on (Costa, Parker, & McCrae, 2000) but only hints about

development in infancy and childhood. There is every reason to believe that children as young as 5 years of age can be meaningfully characterized in terms of the FFM (Kohnstamm, Halverson, Mervielde, & Havill, 1998), but whether 5-year-olds are more extraverted or less conscientious than 12-year-olds is a difficult question. And does it make any sense to even ask about levels of Conscientiousness in infants?

Some readers may be puzzled about how we reconcile the (largely) genetic basis of traits with their lifespan development. One's genes, after all, do not change. But genetic influences are dynamic. They determine development not only in the embryo but in the growing infant and, to some extent, in the aging adult. Parts of the brain continue to grow well into the 20s (Pujol, Vendrell, Junque, Martí-Vilalta, & Capdevila, 1993). Menopause and middle-aged weight gain are programmed into biological clocks. FFT merely adds that personality trait development is also tied to biological clocks: Personality changes can be considered *intrinsic maturation*.

There are three lines of evidence that support this claim, although research is limited so far: First and best established is the cross-cultural evidence we reviewed in Chapter 5. Adolescents differ from adults in the same ways in almost all cultures, despite differences in history, child-rearing practices, customs, and so on. There are other possible explanations for this phenomenon—for example, adolescents all over the world watch television today, whereas their grandparents did not—but these data are certainly consistent with FFT. Development is everywhere the same because it is built into the human species.

Second, McGue et al. (1993) provided a direct test of the above hypothesis by examining the heritability of changes. They followed a sample of twins from age 18 to age 30 and calculated the changes over that period in a series of traits. On average, they saw changes that are consistent with our findings—for example, twins decreased in Stress Reaction (a measure of Neuroticism) and increased in Achievement (a measure of Conscientiousness). But there were also individual differences: Some people decreased markedly in Stress Reaction, whereas others only showed small decreases. McGue and colleagues showed that these changes were heritable: When one identical twin showed big changes, the other twin usually showed big changes as well. These data are consistent with the hypothesis that a genetically programmed biological clock governs trait development. It would be helpful to have replications of this finding, but there are very few longitudinal studies of twins

at this crucial period of the lifespan, where changes are substantial enough to be studied. This ought to be a priority for future twin studies.

The third line of evidence is comparative. King and Figueredo (1997) asked zookeepers from a large group of zoos to rate the personality traits of chimpanzees, using a series of adjectives. Independent ratings by different judges showed reasonably high levels of agreement, suggesting that the ratings reflected something real in the animals' styles of behavior. When the ratings were factored, they found a large dominance factor (well known in studies of animal behavior), but they also found five other factors that resembled the factors in the human FFM. Chimpanzees (at least in zoos) have a long lifespan, often living to 50 years. When King, Landau, and Guggenheim (1998) correlated these factors with age, they found decreases in Neuroticism and increases in Agreeableness (at least among females). It is possible that humans share with their nearest Primate relative a genetic program of personality development.

Why would such a program exist? Conceivably it has evolutionary roots. Adolescents are high in Extraversion and Openness, whereas adults are high in Agreeableness and Conscientiousness. Perhaps at the start of independent life it is advantageous to be adventurous (open and extraverted) to meet potential mates and to learn one's way in the world. Once settled down, Conscientiousness might well be helpful in carrying out the routine tasks of daily life and Agreeableness might be an asset in bonding with children. Interestingly, male chimpanzees are not involved in raising the young and thus (in the present scenario) would have no evolutionary reason to become more agreeable as adults. And, in fact, they do not, according to King et al. (1998).

A few years ago the idea of comparative personality psychology would have seemed ridiculous to many psychologists, but today it is a promising field of research (Gosling, 2001). With the possible exclusion of the Self-Concept, there is no reason why FFT could not be considered a theory of animal personality.

Changes in Individual Differences

Thus far we have been concerned with mean level changes, the normative developmental course of personality traits (e.g., the decline of Neuroticism with age). If everyone changed at exactly the same rate, retest correlations would always be 1.0. But, in fact, we know that retest cor-

relations are never that high, even when corrected for unreliability of measurement. Over very long intervals, like 50 years, only about 60% of the variance is stable (Costa & McCrae, 1992b).

Some writers take that fact as evidence of the effects of life experience on personality. If personality were entirely determined by biological bases, wouldn't it be perfectly stable? Of course, the answer is "no." Consider the twins studied by McGue et al. (1993). As we noted, some showed big decreases in Stress Reaction, some showed small decreases. The effect of that difference would be to shift the rank-ordering of people in the sample and to yield a retest correlation less than 1.0 (in fact, it was .53 for Stress Reaction). We might suppose that the magnitude of the decreases was related to some life experience—perhaps some of them took up meditation or were happily married. But in this case we happen to know that their identical twins showed changes of the same magnitude, although they were unlikely to have experienced the same life events. It is possible that all the changes in rank order are due to individual differences in the hypothesized biological clocks.

That is not the only possibility. Indeed, McGue and colleagues (1993, Table 7) characterized the heritability of changes as modest to moderate in magnitude, leaving room for other causes. Environmental influences are one possibility, but so are other biological causes, such as brain diseases or injuries (even very subtle ones). Longitudinal studies of the brain using magnetic resonance imaging show that there is appreciable atrophy: After age 55, people lose about a teaspoon of gray matter every year (Resnick, Davatzikos, Kraut, & Zonderman, 2000). Depending on where that tissue comes from, it might affect personality traits. With age, human cells deteriorate, and—experience quite aside—the brain of a 70-year-old is not the same as it was 30 years earlier. What is remarkable from a biological point of view is how stable personality traits are.

EVALUATING FIVE-FACTOR THEORY

How should we evaluate FFT as a theory of personality? It certainly summarizes and, in our view, makes sense of a good deal of research on the origin and development of personality traits. It also relates personality traits to many if not most of the other kinds of variables that personality psychologists have studied—attitudes, relationships, identity, and

so on—as well as acknowledging the operation of dynamic processes (such as learning, defense, and long-range planning) that directly generate the flow of experience and behavior. It makes clear statements about how the environment can and cannot affect traits. One could argue that it is very sketchy—there is no catalog of dynamic processes, no theory of identity formation, no principles of psychotherapy to remedy characteristic maladaptations. But the theory is young, and these details can always be filled in later if the basic framework is sound.

But all this probably is not enough to convince most psychologists of the theory's merits. Two objections in particular are likely to be raised: First, the theory does not account for the very traits with which it is centrally concerned. Second, it does not yield novel predictions by which it can be tested. Both of these issues require consideration.

The Structure Postulate asserts that specific personality traits are organized in terms of the five factors of the FFM. Why, you might well ask, are there five factors and not six or seven? And why these factors and not others? It would certainly be possible to create theories. For example, we might try to tie the factors to five distinct regions in the brain, or we might postulate that each solves a particular adaptive problem that led to its evolutionary development. At some point, we might expand FFT to include such postulates.

For the present, however, our view is that the existence of these five factors is simply an empirical fact, like the fact that there are seven continents on the earth (McCrae & John, 1992). If you asked geologists why there are just seven continents, and not five or ten, they would probably say that, at this point in the history of the earth's crust, this is how things have turned out. Millions of years ago there was one continent, Pangaea, that broke apart and, as it happened, we now have seven pieces. Similarly, we would say that in the history of the evolution of the human species (or perhaps the evolution of Primates), the brain grew and developed in such a way as to give rise to the FFM.

This is a crucial point in evaluating FFT. If we had derived the Structure Postulate from a hypothetical mechanism of brain activity or principle of social relations and *then* conducted studies showing that traits really do covary to define the five factors, we would take it as support for our theory—probably as dramatic as the early confirmations of Albert Einstein's prediction that light would bend around massive bodies. But clearly the existing literature is not proof of the Structure Postulate: It is the basis of the postulate. Almost all theories begin with

known facts and try to formulate abstract principles that might account for them. The truth of these accounts has to be tested elsewhere, because many theories would fit any give set of facts; the *right* theory must fit all sets of facts.

We are unaware of any research conducted solely to test FFT, but a great deal of research is relevant to its assessment. FFT was first formulated for an American Psychological Association symposium held in 1992. At that time, there was very little cross-cultural research on the FFM, and cross-cultural findings played no real role in formulating the theory. Subsequently cross-cultural research has blossomed, as we saw in Chapter 5. Do the new findings from this line of research fit with FFT, a model that was based on longitudinal studies of Americans? Indeed they do.

Consider first the invariance of personality structure. As we noted in Chapter 5, Juni (1996) had explicitly doubted that the five factors found in middle-class American samples would be replicated elsewhere in the world. If one holds (as almost everyone did—or does!) that personality is shaped by External Influences, then there is every reason to expect that different traits, or different trait structures, would be found in South Korea, among Black South Africans, and in Lima, Peru. These cultures all differ dramatically from that of the United States in language, in economic and political systems, and in child-rearing practices. Why wouldn't they also differ in personality traits?

The Origin Postulate of FFT states that traits are endogenous basic tendencies, and the Structure Postulate tells us that there are five factors, namely, Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. There is only one human genome, so FFT predicts that these five factors will be found in every culture. We have not yet examined every human culture, but McCrae (2000a) examined 26 cultures and replicated the NEO-PI-R factor structure in all 26 of them. Samples in Korea (Piedmont & Chae, 1997), South Africa (Heuchart, Parker, Stumpf, & Myburgh, 2000), and Peru (Cassaretto, 1999) all show the same structure found in the United States. This is all the more remarkable when it is recalled that the NEO-PI-R had to be translated to be used in these cultures and something might well have been lost in the process.

Again, consider age differences. Cross-sectional studies confound intrinsic maturation with cohort effects and have long been regarded

with suspicion by developmentalists. If birth cohort and the subsequent experience of historical events are important determinants of adult personality, then people who were raised in the People's Republic of China would be expected to show a different cross-sectional profile than people who grew up in North America. Indeed, it is hard to think of a more powerful "natural experiment" than that inflicted by Chairman Mao on his country: the Cultural Revolution. Yet cross-sectional studies of the California Psychological Inventory (CPI) show very similar patterns of age differences in China and the United States, patterns that can be explained by the intrinsic maturation of the five factors (Yang et al., 1998).

How can that be? FFT says that External Influences do not affect Basic Tendencies, which instead follow a species-wide path of intrinsic development. Cross-cultural studies confirm this prediction.

Social Role Theory (Eagley, 1987) holds that personality traits are influenced by the roles assigned to men and women. Because there are huge cultural differences in the treatment of men and women (as recent events in Afghanistan have vividly illustrated), it would not be surprising if there were cultural variation in gender differences in traits. FFT does not offer any guidance about whether there should be gender differences or what they would be, but it does discount cultural influences on traits. Whatever gender differences are found among Americans ought to be found everywhere. Do the data support Social Role Theory or FFT? Clearly, FFT (Costa, Terracciano, & McCrae, 2001). Although the magnitude of gender differences varies, the direction is always the same: Women were higher in Neuroticism, Agreeableness, and Openness to Feelings; Men were higher in Assertiveness and Openness to Ideas in the 26 cultures studied.

Anthropologists and cross-cultural psychologists delight in pointing out the myriad ways in which people from different cultures differ—not only in language and religion, but also in personal space, in kinship patterns, in time orientation, in food preferences, in taboos, in a million particulars. FFT does not deny such differences; instead, it explains them by saying that Characteristic Adaptations are, in large part, the product of External Influences—in this instance, the culture. Undeniable cultural differences can be made compatible with demonstrated universals of personality by separating Characteristic Adaptations from Basic Tendencies, and that is the heart of FFT.

Cross-cultural research is not the only source of support for FFT. Think of the studies of nonhuman animal personality (Gosling, 2001). Because animals have a biology somewhat different from that of humans, the details of FFT would not be expected to be the same, but the general model should be applicable. Any pet owner will confirm that dogs and cats have enduring personality traits, that they form relationships (loving and hostile) with others in their world, that they have habits and preferences, that they act and experience across their whole lives. Animal psychology fits easily into the framework of FFT, but it is hard to imagine a psychoanalytic, or social role, or humanistic theory of animal personality. The fact that the common sense of pet owners is now being scientifically attested (King & Figueredo, 1997; Gosling & John, 1999) can be considered another testimonial to FFT.

It seems likely that progress in molecular genetics and in brain imaging will soon lead to the identification of some of the Biological Bases postulated in FFT, and more direct tests of the theory will then be possible. For example, there are differences in the mean levels of traits across cultures (McCrae, 2000a). FFT asserts that these are biologically based, and genes are the major candidates to explain these bases. Once we have identified the form of a gene (or more likely genes) that underlie Extraversion, we can make a prediction: These forms (or alleles) will be found with higher frequency in cultures that score higher on measures of Extraversion. Such tests are particularly appealing, because genes can be assayed with scientific objectivity and great precision. We need not worry about translations or response sets or the cultural interpretation of items. Of course, personality traits are not genes, they are psychological constructs. But genes may give us a more objective indicator of traits than we have ever had.

* * *

In Chapter 2 we gave a definition of traits as *dimensions of individual differences in tendencies to show consistent patterns of thoughts, feelings, and actions*. That definition applies to populations rather than to people (because people don't have dimensions; instead, they are characterized by their standing on various dimensions), and it pertains to the phenotype, the observable manifestation of traits. We are now in a position to give an alternative definition of traits that applies directly to people and that pertains to the genotype, the underlying causal basis: *Traits are endoge-*

nous Basic Tendencies that give rise to consistent patterns of thoughts, feelings, and actions. From the perspective of FFT, most of this book has been about the developmental course of endogenous Basic Tendencies. But FFT has other components as well. In the final chapter we consider the lifespan development of some other aspects of personality and the influence that traits may have on them.

The Influences of Personality on the Life Course

We have argued throughout this book that personality traits do not change much after age 30. Our conclusion is based not on theories of development or idealizations of aging but on a hard look at such facts as we could find. But when we consider with equal objectivity the work of theorists such as Daniel Levinson and Roger Gould, we cannot escape the conclusion that there is a great deal of change in people's lives. As Gould (1978) wrote, "Adulthood is not a plateau." Most people do not begin a career as soon as they leave school and continue in it uneventfully until retirement. Circumstances change, and people change. Part of the resistance to our message of stability comes from the older adult's recognition that life is often unpredictable and the younger adult's profound hope that there is yet some variety, adventure, and surprise in store. There is.

For some reason, the notion of change without growth seems hard to grasp. Perhaps an analogy would help. By the end of high school, or certainly college, most individuals have reached their peak of intellectual development. They are, in a number of demonstrable ways, not only more knowledgeable but actually more intelligent than they were a few years before. They can think more abstractly, reason more cogently, learn new ideas more quickly, and grasp more intricate patterns. At the

same time there are enormous differences among 20-year-olds in intelligence.

Over the course of the lifespan, we know pretty well what happens to intelligence. Aside from a slow growth in knowledge (as shown by vocabulary test scores) and a progressive decline in reasoning and perceptual ability, particularly late in life, there is rather little change. The average 50-year-old is about as bright as the average 20-year-old, and the above-average 50-year-old is almost certain to have been an above-average 20-year-old.

But this certainly does not mean that the mind is sunk in stagnation for the greater part of life. Novelists use their intelligence to write books, teachers to educate students, business executives to improve management and expand sales. The work of a lifetime does not lead to *development* of intelligence: Nobel Prize laureates would score no higher on IQ tests after years of research than they would have in college. But surely their minds have not been wasted. Personality traits, we submit, are similar in their influence on adult life. Life does not lead to change or growth in personality traits, but it allows a fascinating variety of situations in which personal dispositions, for good or ill, play a part.

WHAT CHANGES?

Let us step back and review for a moment the changes that do occur across the adult portion of the lifespan. Physically, aging brings with it changes in mobility, sensory capacity, strength, and vigor. Certain cognitive functions, particularly perceptual abilities and memory, show declines, gradually at first, but at an increasing rate as we enter old age. The world also changes as we age (with all the headaches that brings for researchers, as we saw in Chapter 4). Our children grow up and move out; our parents grow old and die. Nor are we simply passive victims of the crush of time: We make plans and decisions, change jobs and occasionally spouses. Some of us often and all of us at times think our life is stale and stagnant, but most of us are too busy living to worry about the fact that our personality is not developing.

Even leaving aside External Influences, we must certainly acknowledge that there are extensive changes in a number of aspects of the personality system. Using FFT as a guide, we can briefly review these. Chapter 8 considered the possibility of changes in Dynamic Processes

and concluded that those most central to personality did not seem to show developmental change in adulthood. Of course, there are many, many Dynamic Processes and some of these doubtless change. Cognitive processes like memory and reasoning show well-known changes with age (Salthouse, 1989). There is some evidence that emotional expression becomes muted with age (Bromley, 1978), and there may be some changes in preferred ways of coping or defending in times of stress (Costa, Zonderman, & McCrae, 1991). Because there is as yet no complete catalog, there can be no systematic review of age changes in Dynamic Processes, but that would probably be one fruitful way to approach the topic of adult development.

The Objective Biography

The Objective Biography is the cumulative record of everything a person does and experiences, and it is the one aspect of the personality system that necessarily *must* change with age. Every day lived adds another page to the Objective Biography. There is a sense, however, in which the rate of change decreases with time. As a proportion of life lived, each new day becomes smaller and smaller. The first week of college is full of new experiences; the thousandth week of factory work is pretty much like every other. Still, it is useful to recall that there are always changes of some kind, even in the most routine life structure.

Behaviorally oriented psychologists sometimes attempt to define personality as the sum total of all behaviors. Personality psychologists usually laugh at such a definition (if they don't cry) because it represents an incredibly naïve attempt to maintain the notion that personality encompasses the whole individual while reducing personality to observable facts. In many circumstances behaviors *reflect* personality, although they also, and probably more often, reflect situational demands and constraints and what we have learned by dealing with similar situations. In any case the behavior is not the same thing as the entity it expresses. We would say that the behaviorist's definition reduces personality to the Objective Biography—a peripheral component of the personality system.

Consequently, there is no contradiction at all in saying that personality remains stable whereas behavior changes. All this means is that the situation we face in young adulthood, at a certain point in the history of

the world, is usually quite different from the situation we face in old age. An active, conventionally masculine individual, for example, is likely to be interested in sports all his life. He will probably participate in team and individual sports in high school and college, but once he takes on a family and career his opportunities for active participation in athletics may be cut drastically, since he no longer has the time or the institutional support. As he moves into middle age, physical limitations may pose yet another obstacle to participation. Professional athletes at this point are normally forced to retire or to become coaches or managers instead of players. Amateurs may change their preferred sport, replacing basketball with tennis, or tennis with golf. Physical limitations continue with age, and medical conditions may completely rule out vigorous activity. But an interest in sports and vicarious participation generally continue undiminished: The spirit remains willing, even if the flesh is weakened. And the pace of activity (although not perhaps its vigor) also continues. People who like to keep busy manage to do so in old age by choosing activities that they can handle and perhaps by becoming more efficient in their motions. The range of individual differences even in physical vigor is enormous, and some 80-year-olds keep up a pace that few 20-year-olds can match.

Or consider another example. Physicians trained in the 1950s learned to practice a form of medicine that has since been revolutionized several times. Although many of these men and women are still active today, they certainly do not continue to prescribe the same drugs, use the same diagnostic procedures, or perform surgery in the same way. Their interest in medicine and in dealing with patients has continued unabated over the years, but their behavior has changed dramatically. In fact, it is precisely because of their stable commitment to medicine that they have bothered to learn newer developments.

Characteristic Adaptations

Characteristic Adaptations form the largest and most complex component of the personality system, and although many of them have been studied intensively, it is difficult to make generalizations about what happens with age. Some Characteristic Adaptations are extremely stable: Think of the Southern accent of Senator Strom Thurmond or the pianistic skills of Arthur Rubinstein. Others, like a route habitually

taken to work, change as often as needed to adapt to the requirements of the world. We consider a few of the more important types of Characteristic Adaptations below.

Attitudes and Values

Social psychologists devote much of their time to studying the complex and mysterious processes by which attitudes are changed, but we know without doubt that they do change. New developments—such as nuclear power and genetic engineering—call for new opinions, and at least some people change their minds on old issues as they continue to think about them.

Adolescents and young adults are probably more willing than older adults to adopt new value and attitudes or revise old ones (although it is not clear that there is much difference between 30- and 80-year-olds in this regard). There are two major reasons for this. One is that as a person ages, genuinely new ideas become increasingly infrequent, and, as R. W. White (1952) remarks, “accumulated experience . . . more and more outweighs the impact of new events” (p. 333). Another is that, once settled into a life structure, we have much more to lose by changing basic values. The executive rising on the corporate ladder has little incentive to embrace socialism; the lawyer who has invested years in learning her profession is understandably reluctant to question it. Our values, like our personality traits, help to shape our life structure and are themselves perpetuated by it.

But people do change their minds, just as they sometimes change religions, political parties, and occupations. In fact, the enduring personality disposition of Openness to Experience is in part characterized by the ability to keep an open mind, to consider new opinions, and—at least occasionally—to change attitudes and values (McCrae, 1996). We would expect open people to revise their ideas and values repeatedly, whereas closed people would cling tenaciously to the opinions of their parents and other respected authorities.

Note, however, that change is not necessarily growth. Openness to Experience does not guarantee that new opinions and attitudes will always be better, wiser, more differentiated, or more in keeping with contemporary reality. Some open people are simply flighty, moving from one world view to another with an amazing regularity. When he or she combines Openness with critical capacity, however, the open person

certainly appears to be in a better position to adapt to a changing world. The point is that change in opinions is not only consistent with stability of personality: It is itself an enduring quality of some individuals.

Social Roles

Few ideas in the social sciences can match the scope or power of the concept of *role*. Social-psychological jargon is full of qualifications of the basic metaphor: role expectation, role performance, public role, deviant role, even roleless role. Role theorists (Goffman, 1959; G. H. Mead, 1934), who tend to be sociologists rather than psychologists, have spent considerable time and effort in proposing and debating various definitions of that eminently useful concept, but for our purposes most of them are equally good. Allport (1961), for example, defined a role as “a structured mode of participation in social life. More simply, it is what society expects of an individual occupying a given position in the group” (p. 181).

Roles are so peculiarly useful because they bridge the gap between the individual and the society. On the one hand, roles are functional units of social systems, and without even mentioning individual persons a sociologist can describe a society by specifying the nature and interaction of roles. The traditional nuclear family, for example, consists of a father and mother (in the complementary roles of husband and wife) and one or more children, who, in addition to being sons and daughters to their parents, are also sisters and brothers to each other. Everyone in this system has a predefined part to play: Parents are supposed to take care of the children but also to control and guide them; children are supposed to love and obey their parents. (The fact that this model family is so seldom seen in reality is a cause for distress to all concerned, but a sophisticated role theorist would point out that the role of teenager in modern America is defined as much by defiance as by obedience.) Other social systems are also understandable in terms of roles: A business consists of bosses, workers, and customers; a city has officials and citizens.

On the other hand, roles serve an equally important function for individuals. As Characteristic Adaptations, they can be considered scripts that tell people what to do in specific situations. For example, a woman who has internalized the role of *supervisor* has learned the skills needed to manage others and understands the obligation to make work

assignments, evaluate performance, train new employees, and so on. Much of her behavior—at least during business hours—can be attributed to her role as supervisor.

Some psychologists take the view that personality is, or is largely, the collection of roles one plays. To one who maintains such a conception, the contention that personality is stable must seem somewhat silly. Don't people become parents and then grandparents? Don't they retire or make midcareer shifts? Doesn't age increase the probability that one will adopt the sick role? Of course. If one adopts a role theory of personality, the whole issue of stability becomes trivial.

Other psychologists, also influenced by social-role theories, distinguish between roles and a deeper layer of personality, but they expect that acting out certain roles will lead to changes in at that deeper level. The man who takes a job as a policeman may become authoritarian, so the theory goes, because everyone thinks he is and treats him accordingly (G. H. Mead, 1934). In laboratory settings, this kind of manipulation can induce changes in people's reports of what they are like, and social psychologists have sometimes asserted that age-related "role transitions . . . must continuously redefine our personalities" (Veroff, 1983, p. 341), especially in old age (M. M. Baltes & Schmid, 1987).

Our findings of stability present a problem for this theory. Assuming that our measures are taken seriously, two possibilities emerge as likely explanations. The first is that because role changes are not as pronounced as we might have thought, they leave personality basically unaltered. Perhaps the change from mother to grandmother is not very significant in comparison with differences between fathers and mothers. Again, retirement from an occupation may not be a radical change if, as Havighurst, McDonald, Maculen, and Mazel (1979) have shown, many people continue their professional activities after formal retirement. Age roles do not seem to be a major influence on social identity as seen in the spontaneous self-concept (McCrae & Costa, 1988a). The second possible interpretation is that this role-based theory of personality influences is simply wrong. Perhaps people do not internalize the perceptions others have of them, laboratory studies notwithstanding. Swann and Hill (1982) showed that subjects given feedback inconsistent with their own conceptions of their personality traits changed their images of themselves—but only until they were given a chance to reassert themselves. By distinguishing between Characteristic Adaptations and Basic

Tendencies, FFT can explain why changes in roles need not affect personality traits.

Interpersonal Relationships

Some aspects of interpersonal relationships are based on roles, and as roles change, so do relationships. Soldiers who fought on opposite sides can be friends when the war is over; students can become the colleagues of their professors when they graduate; retired executives can no longer dominate others as they used to. These changes in relationships are perhaps the most significant feature of changes in roles.

Other aspects of interpersonal behavior are expressions of traits that cut across many different roles and impart a personal flavor to their interpretation. As far as role requirements allow, the friendly person is likely to respond with warmth to employers, neighbors, and relatives alike. The shy person will be self-conscious among strangers, but also with friends when he or she is too much the center of attention. Because these dispositions cut across roles, the changes in roles that come with age do not affect their expression and they tend to remain stable across the adult years.

But there are some relationships that are shaped by more than social-role requirements or personal dispositions. Farrell and Rosenberg (1981), for example, in their discussion of husbands and wives at midlife, acknowledge that the adage that opposites attract has rarely been supported by research but point out that couples seem to develop a pattern of complementary behavior. Two assertive individuals may make a couple in which one is clearly dominant, or they may carve out separate domains in which to exercise their dominance (traditionally, the wife ran the house and the husband made decisions about the car). Whenever two individuals share a significant and lasting relationship, its characteristics are likely to be the result of the particular ecology of both of the persons and of the situation, a series of compromises and adaptations evolved over a period of time.

We might describe a relationship in terms of patterns of power and dependency, the emotional support given or received, or the depth and nature of the bond of love or hatred. Little is known about the stability of these features of relationships. Do newlyweds quickly set up a pattern that will endure throughout married life, or are the terms con-

stantly renegotiated as new problems and opportunities emerge—the birth of a child, a career for the wife, an older parent moving in? Is it possible that there are regular developmental changes that characterize relationships, even if there are no developmental changes in individuals' traits? Do wives take over the decision-making roles in the family after middle age even though they do not otherwise become more assertive or masculine?

One set of relationships does show change closely related to age. As we age, we become independent of our parents and our children become independent of us. Interview studies seem unanimously to suggest that the intrapsychic rate of change by no means keeps pace with the behavioral. Years after moving away from home and taking financial responsibility for themselves, adults sometimes seem curiously tied to the approval and opinions of their parents. Some writers argue that it is only the shock of a parent's death or chronic disability that convinces middle-aged men or women that they are really as big and strong as their parents once were. The struggles of the adolescent to break away from parental domination sometimes result in periods of many years in which there seems to be no feeling at all left for parents. Levinson's interviews suggest that at middle age the feelings, positive and negative, may reawaken and lead to a new, perhaps more mature relationship. Is there a developmental progression from individuation in which young adults emphasize their separateness from their parents, through a period of increasingly better perspective on the parents as individuals, to full peer-like mutuality between adult children and their parents?

It is of some interest to note that studies of the external trappings of family life show great continuity between generations (Troll, Miller, & Atchley, 1979). The majority of adult children see their parents at least once a week, and considerable mutual support is offered in the form of money, favors, and so on. These family ties generally continue until death; the notion that their children abandon most old people is largely myth.

The Self-Concept

The Self-Concept is technically a part of Characteristic Adaptations, but it is a special part through which we understand ourselves. Our view of our own traits is a part of our self-concept, and, to the extent that traits are stable, so is the Self-Concept (Mortimer et al., 1981). But roles and

relationships also play a major role in our self-definitions; as they tend to change over time, so must the Self-Concept. As we saw in Chapter 8 in the study of the spontaneous self-concept, when people are asked to respond to the question “Who am I?” they frequently describe themselves by reference to what is called their *social identity*: I am a doctor, a Republican, a mother, a Lutheran. Each person occupies, simultaneously and successively, a dazzling variety of roles, and one of the chief functions of the ego described in Chapter 8 is keeping all of these straight. Much of the individual’s sense of identity is tied to the set of roles to which he or she is more or less permanently committed. But much of what might reasonably be called adult development is an adaptation of the Self-Concept to changing roles.

The word *identity* is sometimes used to refer to those aspects of the Self-Concept that are essential to one’s self-definition. Personality traits may be a salient part of one’s identity, but so may social definitions or body image. Whether and in what ways identity changes in adulthood are only beginning to be explored. Whitbourne and Waterman (1979) have traced the development of identity from college age (by which time Erikson suggested it should be formed) through the 30s, and they report an increase in commitment over this period—a conclusion that echoes R. W. White’s (1952) finding of “stabilization of ego identity” in his case studies covering a similar period of development. Does identity continue to stabilize, or are the identity crises of adolescence revived at middle age? Our studies (McCrae & Costa, 1988a) and some others (e.g., Whitbourne, 1986a) point to stabilization in the Self-Concept and perhaps also in those central elements that constitute one’s identity.

Another aspect of the Self-Concept is the life narrative (McAdams, 1993). According to proponents of this construct, people do not organize their Self-Concept alphabetically, keeping lists of relevant traits, roles, and relationships. Instead, they structure it in terms of a story, intended to give meaning, purpose, and unity to their lives. Some see their lives as an unfolding tragedy; some, as a comedy. Some focus on the obstacles they have faced and overcome; some, on the friends they have made. Most people base their life stories on their Objective Biography, but they do so very selectively.

The Objective Biography changes with each passing day. It seems unlikely that the life narrative would change so rapidly. Major life events would certainly offer an opportunity for reconstruing one’s life, but because there have been no longitudinal studies of life narratives,

we do not yet know whether or how often that happens. It seems likely that some aspects of the life narrative—like optimistic or pessimistic tone—are tied to personality traits and thus might be expected to be stable. Other aspects, such as the choice of critical events, probably change.

STUDYING LIFE STRUCTURE AND LIFE COURSE

In view of the evidence of trait stability in adulthood, there would seem to be two directions in which the study of adult development might proceed: It might focus on samples where substantial change would be expected, including individuals undergoing psychotherapy or those under age 30; or it might redefine the field as the psychology of adulthood, in which enduring dispositions would be viewed as a foundation for adaptation to the changing demands of life.

We believe both of these approaches are useful: Researchers should study change where it occurs and stability where it does not. There has been a renewed interest in personality changes during the college years (R. W. Robins et al., 2001; Neyer & Asendorpf, 2001), studies that can profitably be extended at least through the decade of the 20s. (Studies of personality change prior to college are surprisingly few; they suggest that there is considerable individual change but relatively little mean level change between 12 and 18; Costa, Parker, & McCrae, 2000).

Psychotherapy remains an area in which much remains to be done. Longitudinal studies are observational, not experimental; they tell us what happens to a sample of people during the normal course of events. The fact that most people do not change much does not necessarily mean that they could not change given the right set of circumstances. Although interest in psychotherapy research waxes and wanes (VandenBos, 1986), we still do not have long-term, large-scale studies of the effects of different forms of therapy on personality traits.

Equally exciting to us is the research agenda that flows from a reversal of the conventional approach to adult development. Students of age and personality have traditionally asked, "How does aging affect personality?" The complementary question—"How does personality affect the aging process?"—is still new. In the first version of this book (McCrae & Costa, 1984), we lamented the scarcity of research on this topic. In the intervening years, things have begun to change (Costa &

McCrae, 1989b). There is a Life History Research Society which considers such topics as "Straight and Devious Pathways from Childhood to Adulthood" and "Pathways to Adult Competence," and Lachman (2001) has edited a *Handbook of Midlife Development* in which the role of personality traits is occasionally discussed.

Much remains to be done in charting the influence of each of the five domains of personality on people's careers, familial relationships, and adaptation to stress and disease, but enough has already been done to show the powerful effect personality traits can have. For example, Conley (1985) traced the couples who had participated in E. Lowell Kelly's pioneering longitudinal study (1955)—a study that had provided some of the first strong evidence of personality stability. Conley used self-reports and ratings gathered in 1935 and 1955 to predict aspects of life in 1980 and found evidence of pervasive influences of traits. Men and women rated by peers in 1935 as being high in Neuroticism were more likely to have had an emotional disorder in the succeeding years; men rated as poor in impulse control (or Conscientiousness) were more likely to have become alcoholics.

Conley sampled a number of variables that give a flavor of daily life in 1980 for these men and women. Those he called intuitive types, whom we would describe as open to experience, had artistic hobbies and preferred to watch public television. His sensing types, closed to experience, listed gardening, sewing, and cooking as hobbies and preferred to watch game shows. Dispositions assessed early in life can make sense of the activities of older individuals.

A recent study by Soldz and Vaillant (1999) traced Harvard men from college to old age. In their last year of college, the men were rated by a psychiatrist on a set of 25 traits, which were later scored in terms of the FFM. The men completed the NEO-PI in 1988. Despite the difference in instrument and source of data (psychiatrist ratings vs. self-reports) and the lapse of over 40 years, there were significant correlations for Neuroticism, Extraversion, and Openness (r 's = .19 to .38). Of more interest here, however, is the effect of these stable traits on the life course. Soldz and Vaillant had access to a variety of outcome measures collected at various points between the two assessments. All the college ratings were related to at least one outcome: Those high in Neuroticism smoked more; those high in Extraversion earned more money. Open students showed more evidence of creative productions and were less likely to have conservative political attitudes; they also were more likely

to have consulted a psychiatrist at some point. Men rated as being agreeable in college showed more effective social relationships in middle age. Men rated as being conscientious showed less alcohol abuse and less need of psychiatric consultation, and reached the highest levels on an Eriksonian measure of adult development.

The most fundamental influence of personality stability on life is the fact of stability itself. As we noted in Chapter 7, in order to make a life for oneself at all one must have some firm foundations, some predictable regularities. Some sources of continuity are external to us, including our cultural background and social class and socialization into particular professions or groups. Others are internal, including talents, abilities, and personality traits. Choosing a major in college, deciding whom to marry, accepting a job transfer, planning for retirement—all depend on the knowledge that our basic motives and styles will endure. Bandura (1982) argued that part of the reason for stability in personality may be that we choose and create environments that will reinforce our dispositions; regardless of the reasons, stability is a fact we must and do count on.

Psychological Adjustment across the Lifespan

One of the saddest truths that has emerged from the experience of psychologists in the last century is that individuals with difficulties in living generally continue to have problems, even after extensive treatment. Tsuang, Woolson, Winokur, and Crowe (1981) have shown that 92% of individuals once diagnosed as schizophrenic are still judged schizophrenic 30 years later. Twenty-five years after World War II, Keehn, Goldberg, and Beebe (1974) found that soldiers given psychiatric discharges were more likely to have died from suicide, homicide, accidents, or alcohol-related illnesses than normal soldiers with the same war experiences. L. N. Robins (1966) showed that deviant children are much more likely than others to become adult criminals, and recidivism rates from police records show that an adult criminal generally leads an entire lifetime of crime. But schizophrenia is an illness, and the social system may trap some individuals into a pattern of crime. Do enduring personality traits in normal individuals affect long-term adjustment to life?

Certainly. Neuroticism is the aspect of personality most relevant to adjustment, and those high on this dimension are likely to show evi-

dence of maladjustment at all ages. They are, for example, likely to be dissatisfied with life and exhibit low morale (Costa & McCrae, 1980a). Maas and Kuypers (1974) found in a 40-year follow-up study of a group of men that "old age does not usher in or introduce decremental psychological processes. Rather, old age may demonstrate, in perhaps exacerbated form, problems that have long-term antecedents" (p. 203). There is evidence that, regardless of age, individuals high in Neuroticism are more likely to use ineffective coping mechanisms such as hostile reactions, passivity, wishful thinking, and self-blame in dealing with stress (McCrae & Costa, 1986). Vaillant (1977) has documented the pervasive influence of neurotic coping styles on the lives of an elite group, Harvard graduates.

The antecedents of adult Neuroticism can already be seen in childhood, and its lifelong consequences have begun to be traced. For example, children with a history of temper tantrums (who were perhaps low in Agreeableness as well as high in Neuroticism) later showed disturbances in career and marriage, with a higher probability of downward occupational mobility and divorce (Caspi, Elder, & Bem, 1987). Those veterans who were characterized by higher levels of inadequacy and anxiety during adolescence were more likely to have had stress reactions after World War II than were other combat veterans (Elder & Clipp, 1988).

Liker and Elder (1983) set out to explore the joint impact of an environmental event (downward economic mobility caused by the Great Depression) and the disposition of Neuroticism on a significant life outcome: the quality of marital relations. They found that the emotional stability of both partners had a continuing impact on the quality of their marriage, although the quality of their marriage had no influence on their personality.

In a sense, this is nothing new; it is merely a restatement of the fact that Neuroticism is an enduring and consequential disposition. It is in interaction with age and life circumstances that the influence of Neuroticism becomes most intriguing. Consider, for example, individuals who experience a so-called midlife crisis. Everyone acknowledges that there are such people, although there is still some dispute as to how many. The problems and concerns of this syndrome are distinct: They revolve around the themes of lost youth, abandoned dreams, and meaningless relationships. When Farrell and Rosenberg (1981) interviewed five subjects who seemed to be in this state, they concluded from detailed life

histories that this was not simply a “developmental emergent” without precedent in the life of the individual. Even if the individual had led a relatively solid and satisfying life up to that point, these researchers believed that the seeds of crisis had already been laid by defects in the subject’s character and the resulting poor choices he had made.

We found evidence of another sort for the same phenomenon. In our questionnaire studies of the midlife crisis (described in Chapter 9), we found that only a minority of men of any age felt themselves to be in a crisis. When, however, we examined the scores these few had received on personality measures taken 10 years earlier, we found that even then they had been significantly higher than others in Neuroticism. If we spoke to these men today, we might find them blaming all their troubles on recent events or on their age or health. But they had had more than their share of complaints many years earlier, and it begins to seem that they carried their troubles with them. It may be that the *form* of the trouble varies with the period of life in which they happen to be. Gould (1972) found that the nature of the concerns that brought patients to therapy differed by age: Adolescents had problems with their parents; young adults, with career choices. Well-adjusted individuals who go through the same situations regard them as challenges rather than insoluble problems.

Such studies as these point the way to whole new lines of research, since they show that neither our dispositions in themselves nor the circumstances that age and history confront us with are sufficient to explain the forms (or failures) of our adjustment. Instead of being content to ask “How do older people adjust to retirement or bereavement or relocation?” we may find that we have to ask “How do older neurotic people . . . ?” or “How do older open extraverts . . . ?” We need to develop a contextual psychology of aging that integrates personality traits with changing life situations and social roles.

Personal Projects, Social Clocks, and Psychobiography

Much of the conduct of life is routine: going to work, doing errands, cleaning the house, and watching television. However humdrum they may seem, these routines are generally the result of an elaborate evolution intended to allow us to gratify most of our needs most of the time as well as we can in our circumstances. What seems unremarkable from its familiarity is in fact an intricate construction, a life structure. As we

continue to age a succession of life structures forms a life course, our own personal history.

As a technique for studying the structure of life at the day-to-day level, Little (1983) has proposed a new unit of analysis, the *personal project*. Personal projects are sets of acts intended to achieve a goal; they range from "getting a better job" to "overcoming my shyness" to "going to the ballet" (p. 293). Individuals appear to be able to list the projects that concern them at any given time quite easily and can then rate them on such dimensions as importance, difficulty, control, and absorption (the degrees to which they are involved in and preoccupied with them). One advantage of personal projects as units of analysis is that they provide a detailed specification of the life structure at a single time, one that is phenomenologically real to the individual. An occupational history that includes only type of work and dates of hiring, promotion, and firing misses the richness that can be drawn from an analysis of personal projects: Are job concerns central to life, or simply an economic necessity? Are tasks part of an ongoing career plan or a series of unrelated activities? Is the work pleasant or unpleasant, fascinating or dull, imposed or chosen?

Personal projects are also intriguing because the dimensions on which they are rated can be related to dimensions of personality. Little, Lecci, and Watkinson (1992) reported that highly conscientious people feel their projects are progressing well, that open people believe their projects are particularly meaningful, and that those high in Neuroticism find their project stressful. These kinds of analyses help to spell out the implications of personality traits and their influence on life at a molecular level.

On a larger scale, Helson, Mitchell, and Moane (1984) examined effects of personality on *social clock projects*. Some life goals, such as marrying, raising a family, and having a successful career, are shared by nearly everyone in the culture and have a prescribed time course: There is a social clock by which we can tell if our life is progressing at the rate it should. Helson and her colleagues examined personality scores gathered at the time of college graduation (between 1958 and 1960) for a group of women they subsequently followed. Nearly all the women of this group professed a desire to marry and start a family. They found that "women who started their families on time may be characterized as having been confident and assertive in college in addition to having the motivations and skills to adjust in conventional

ways" (p. 1083). We would probably describe them as conscientious extraverts. A minority of women in this sample aspired to a career in addition to marriage. Those who succeeded by age 42 had been assertive, forceful, self-confident, and ambitious in college; they appeared to be more open and extraverted than the women whose career plans failed.

Although society may dictate the major directions of life, individuals in modern societies have considerable latitude for personal choices, some of which reshape the life course. These critical choices, and the reasons behind them, form the substance of biographies (Herold, 1963): Napoleon Bonaparte decides to restore order to the French Revolution, in part because of his conservative social views; he has himself crowned Emperor to satisfy his quest for personal aggrandizement; he divorces Josephine and marries Archduchess Maria Louise (daughter of the Austrian emperor) to ensure the rights of his progeny; he invades Russia because he believes he can bully Czar Alexander I into submission; he returns from Elba because he has become convinced that, despite all odds, it is his destiny to rule. Biographers generally feel that the best way to understand the individual's personality is to record the subject's life, to analyze the circumstances that helped shape the life course (in part to separate external influences from the inner characteristics), and then to interpret the action in terms of the inner drives, talents, or weaknesses of the person (Runyan, 1981). Erikson (e.g., 1962) devoted much of his career to such psychobiographies, and most clinicians rely heavily on the life history as a means of understanding their clients.

In Chapter 9 we argued that life histories should not be used as the basis of a psychology of personality development; rather, we think that knowledge about personality and its continuity in adulthood can form a basis for understanding life histories (see Nasby & Read, 1997, for a case study interpreted in terms of the FFM). The psychodynamic approaches of Freud and Erikson have had a tremendous influence on the thinking of historians and literary critics as well as psychologists, because psychoanalytic formulations can be used to make sense of historical and literary figures. The same benefit can be offered by trait psychology; in fact, the stronger empirical basis of trait psychology makes it, in our view, a much better candidate for this role. Historians, of course, generally cannot administer standardized personality questionnaires to their cases, but they can use the behavior patterns of a lifetime to estimate standing on dimensions of personality and then use these as a

framework for interpreting the significant events in the lives of their subjects (see Winter & Carlson, 1988).

It is possible to gather personality ratings on historical or literary characters from panels of experts using standard questionnaires such as the NEO-PI. These ratings provide a more objective alternative to the subjective interpretations usually offered by historians and literary critics. For example, Rubenzer, Faschingbauer, and Ones (2000) asked historians to rate each U.S. President on the NEO-PI-R and averaged the results to gain consensus judgments. They found evidence of agreement across judges: George Washington was high in Conscientiousness; Abraham Lincoln, in Openness. On average, Presidents were closed, but the Presidents considered greatest (e.g., Lincoln, Theodore Roosevelt) were open. Richard Nixon was profoundly disagreeable (McCrae, Yang, et al., 2001), but he was high enough in Conscientiousness to be twice elected President. (For an example of how personality research and literary criticism can fruitfully interact, see McCrae's [1994] analysis of Herman Hesse's *Narcissus and Goldmund*.)

Empirically oriented psychologists have often felt uncomfortable with psychobiographies, and with good reason: Various explanations can usually be offered for the same facts, and there is no conclusive way to choose among them. Different historians have vastly different views of the character of Napoleon and make widely different hypotheses about what he would have done after his return from Elba if the Duke of Wellington had left him any choice. But we cannot conduct an experiment, cannot go back in time and manipulate history so as to test these theories. And, according to most philosophers of science, what is untestable is beyond the bounds of science.

However, as Birren and Hedlund (1987) pointed out, it is perfectly reasonable to examine biographies as a source of testable hypotheses. If we are interested in people in general, rather than in some particular historical figure, we can move from speculative biography to testable psychology. We can also abandon the post hoc method of looking for traits that underlie life choices and begin a systematic study of the effects of specified traits on individual lives. We are not all Napoleons, but we all make choices: if and when and whom to marry; whether to have children and how many; what career to pursue, when to change it, when to persist despite adversity; whether to move to another city and reestablish friends and routines. Curiously, the two most significant and most thoroughly researched life choices—marriage and career—show quite different patterns.

Marriage and Divorce

Theories of marital choice have generally looked for a basis in either similarity or complementarity. There is good evidence that people do marry others of similar social, ethnic, and religious backgrounds and that physical attractiveness is a powerful determinant—the rich and beautiful tend to marry the rich and beautiful. There is also clinical evidence, of the kind reported by Farrell and Rosenberg (1981), that marital relationships often evolve complementary functions. But when personality measures are used, they frequently fail to predict marital choice: Extraverts are just as likely to marry introverts as they are to marry extraverts; adjusted people choose neurotic spouses as often as adjusted ones. There seems to be a small influence of similarity in the domain of Openness—open people tend a bit toward marrying open people (McCrae, 1996)—but that may in part be an artifact of similarities in social class, education, and so on: When two college graduates marry, both of them are likely to be more open than high school drop-outs.

This might seem to be a defeat for trait psychology: The major dimensions of individual difference seem to have little influence on one of the most significant decisions of our life. Indeed, it is a healthy reminder that traits cannot explain *everything*. But it leads to a series of fascinating and thus far almost completely unexplored questions. Why would a well-adjusted person take on the difficult task of trying to live with an anxious, hostile, and depressed mate? Why would an open person consent to live with one who was closed? And, more than why, how do they manage together? When an introvert marries an extravert, does it lead to conflicts in their joint social life, do they arrange to live separate lives, or does the extravert “bring out” the extraverted side of his or her spouse? Does it matter whether the husband or the wife is the introvert? Will the husband’s preferred style of socializing prevail regardless of what it is? Psychologists have spent a great deal of time studying how one individual adapts to an environment, but how do two individuals adapt to the environment they create for each other?

Buss and Barnes (1986) have begun to study some of these issues by looking at mate preferences. Individuals seek spouses with different characteristics: kindness, intelligence, attractiveness, wealth, and so on. In Buss and Barnes’s study these preferences tended to reflect personality characteristics—although there appeared to be frequent sex differ-

ences. For example, kind and considerate spouses were preferred by men high in Extraversion and by women high in Neuroticism. Preferences, in turn, were related to the actual personalities of spouses: Women who preferred kind husbands tended to marry men who scored high on measures of Agreeableness. It appears that personality is related to marital choice, but not in a straightforward fashion.

There is considerably more information on divorce. Not surprisingly, marital dissatisfaction and divorce are chiefly related to Neuroticism, as a number of prospective longitudinal studies have shown. Between 1935 and 1938 E. Lowell Kelly began a study of 300 engaged couples. He obtained personality ratings on both partners from five acquaintances; 45 years later he assessed marital status and satisfaction (Kelly & Conley, 1987). Neuroticism in both husband and wife and low Conscientiousness in the husband predicted divorce. An analysis of the causes of divorce suggested an explanation:

A wide variety of reasons for divorce were reported, but a number of these (in particular, emotional instability, arguments, emotional overreactions, and sexual problems) are probable manifestations of Neuroticism, and others (particularly infidelity, alcohol abuse, and irresponsible social behavior) are probable manifestations of low impulse control [or Conscientiousness]. (p. 32)

Kelly and Conley also considered the case of individuals who remained unhappily married. These men and women had been rated as being higher in Neuroticism, but the men were also seen as introverted and disagreeable. Perhaps aloof and antagonistic men saw little prospect of a better second marriage and settled for what they had.

Interestingly, Kelly and Conley had measures of many other variables thought to be important to marital adjustment: early home environment of the partners, including their parents' marital adjustment; attitudes toward marriage during the engagement period; sexual histories; and stressful life events during the marriage. All these variables were far less important than personality in predicting marital adjustment.

Careers

On the second issue, career choice, the influence of personality is both better understood and more obvious. Given a choice, people tend to-

ward occupations that allow the expression of their personality traits. This is most clearly seen in inventories of vocational interest, such as the Holland (1985) Self-Directed Search, which provides scores for six categories of interests. Individuals with enterprising and social vocational interests tend to be extraverts; those with artistic interests tend to be open. Investigative interests are found among intelligent introverts, and conventional interests among those closed to experience. Not surprisingly, open people report interest in a wider variety of occupations of all sorts (Costa et al., 1984).

Similar relationships are found when other vocational interest inventories, with different categorizations of interests, are used. And after age 25 vocational interests are known to be extremely stable. Some of the best early evidence for stability in personality was provided by Strong's (1955) 25-year retest of vocational interests.

When we move from vocational interests to actual occupation, the data become less clear. Open people may be interested in the occupation of poet or concert pianist, but how many of them have the talent or the financial resources to make a career of these interests? At this point in history at least, one's gender and educational level are the primary determinants of occupational choice and economics is the major incentive. By and large, people take the highest-paying job they are qualified for, regardless of their preferences. Given a choice, they will gravitate toward the positions best fitted to their temperament—but most of us are not really given a choice. Nevertheless, studies do show that salespeople are more likely to be extraverts than introverts, and that cooks are less open to experience than are creative writers (Cattell et al., 1970). Data from the Myers-Briggs Type Indicator (MBTI) suggest that most psychologists are high in Intuition, or what we would call Openness; those who are also high in Agreeableness are more likely to be clinicians, whereas those low in Agreeableness gravitate toward experimental psychology (Myers & McCaulley, 1985).

We know from studies in Boston (Costa & McCrae, 1978) and in Baltimore (McCrae & Costa, 1985a) that individuals who have started a new career or switched to a different line of work are more likely to be open than closed to experience. This is as true for women as it is for men. Longitudinal studies have also allowed us to separate cause from effect: People are not open *because* they were broadened by a new vocational experience; instead, they were open to experience years before

they made the change, and so it seems likely that Openness contributed to the decision to start a new career.

Ordinarily, people do not just sign up for the job they want. Instead, they have to find a job and convince their prospective employer that they should be hired. The job search is often a long and potentially frustrating one. Are personality traits involved? A recent review (Kanfer, Wanberg, & Kantrowitz, 2001) concluded that they are indeed and that the most important traits are related to Extraversion. Extraverts, it appears, really are go-getters, willing and able to sell themselves. Perhaps that is why, in the long run, they end up making the most money (Soldz & Vaillant, 1999).

Once in a job, personality is likely to manifest itself in the way the job is performed (Hoekstra, 1993). The extraverted bank clerk will spend more time than others chatting with customers; the librarian who is closed to experience may make more efforts to maintain quiet than to encourage reading. And if the gap between the occupational demands and personal disposition is too large, the individual is likely to quit and find a different line of work.

Between 1970 and 1990, most industrial/organizational (I/O) psychologists ignored personality traits on the assumption that traits did not predict job performance. But in 1991 Barrick and Mount published a large-scale review in which they classified the many personality scales that had been used in research according to the FFM. When organized in this way, a clear pattern emerged: Job performance in almost every category was predicted by Conscientiousness. Employees who are well-organized, punctual, and task oriented do well, whether they are managing corporations or digging ditches. That simple and, in retrospect, rather obvious truth came as a revelation to I/O psychologists and led to the prompt rehabilitation of trait psychology. Other traits also affect job performance, but they tend to be specific to certain kinds of jobs. Low Neuroticism is important for police (Costa, McCrae, & Kay, 1995); Agreeableness is valuable in the service industry (Costa & McCrae, 1995).

TRAIT INFLUENCES ON THE SELF-CONCEPT

To this point we have been concerned with how personality traits affect the Objective Biography, seen broadly as the life structure and life

course. We can also ask whether traits affect the internal reflection of these elements in the Self-Concept: identity and the life narrative. In one sense, as we have seen before, traits must affect the Self-Concept, because they are among its most salient contents. Extraverts think of themselves as extraverts, and they will tell you so if you ask them to answer the question "Who am I?" (McCrae & Costa, 1988a).

However, at present we are concerned with a more subtle question: Do traits affect the development of the Self-Concept over time? Do open individuals have a different way of thinking about themselves, and about changes in themselves, than closed individuals? Does Neuroticism bias the selection of events one chooses to discuss in one's life narrative? These questions have seldom been asked, but they certainly merit consideration.

Identity

By *identity* we mean those aspects of the Self-Concept that are particularly salient or definitional—characteristics which, if they were taken from us, would lead us at least briefly to wonder if we were still ourselves. Traits form part of many people's identity, but so do values, roles, and relationships. Because these latter Characteristic Adaptations change over the life course, it becomes a matter of interest to see how the Self-Concept changes and whether the changes are moderated by personality traits.

It is useful to begin by recalling that the Self-Concept is not a passive mirror image of one's Characteristic Adaptations and Objective Biography. There is no one-to-one correspondence between the roles one plays in society and the definitions one has of oneself. The fact that one works from 9:00 to 5:00 as a secretary does not mean that being a secretary is part of one's self-definition. In fact, many secretaries probably see themselves as homemakers, novelists, or actors who are merely taking a job as a secretary; their dreams and aspirations are elsewhere. Conversely, even when we fill a role that we have always hoped to fill, we may not really believe we fit in. In the eyes of the law, we may be fully adult at age 18 or 21; in our own eyes it may take many more years till we reach adulthood. In the meantime we act as adult as possible and hope that others will believe our bluff. Symbols are often important in establishing our sense of identification with a role. Stethoscopes and white coats may be more important than licenses in making men and

women feel like doctors; military and religious uniforms are indispensable in creating the sense of commitment and identification that is required by these institutions.

A number of clinicians have remarked on the discrepancies between the face we present to the world—our persona or mask, as Jung called it—and our real sense of self. Alienation and a sense of meaninglessness may result from acting out roles with which we cannot identify. Conversely, we may identify too closely with superficial roles at the risk of being at best shallow, at worst seriously out of touch with our own inner needs.

Why do we identify with some roles and not others? Sheldon, Ryan, Rawsthorne, and Ilardi (1997) showed that traits are one determinant. They asked people how they behaved in various circumstances and how they felt about themselves in those roles. They found that people did vary somewhat in their behavior but maintained a kind of loyalty to their basic traits: People felt happiest and most genuine when they acted in accordance with their enduring dispositions. An introvert who needs the money may take a job as a salesman, for example, but he may never really feel he belongs there.

The idea that people identify with roles that are consistent with their personality traits has a corollary, namely, that people who readily adopt a variety of roles are likely to have a shaky and uncomfortable identity. This view runs counter to other models of adjustment that suggest that role flexibility ought to be adaptive. After all, people must wear many different hats and act in somewhat different ways for each role; wouldn't it make sense to argue that a highly differentiated self-concept would lead to greater psychological adjustment?

Donahue, Robins, Roberts, and John (1993) tested these alternative hypotheses by giving a measure of self-concept differentiation to a sample on whom longitudinal observations had been made over many years. What they found was that more differentiated people were less well-adjusted and had been for many years. It seems likely that Neuroticism interferes with the development of an integrated self-concept. It is perhaps not surprising that individuals high in Neuroticism, whose authentic nature is to be miserable, keep trying new self-definitions, like an insomniac who cannot find a comfortable position in bed.

Openness to experience may also figure in the correspondence between inner and outer versions of the self, and provide a sense in which flexibility is adaptive. Those who are open to their own feelings, willing

to think out and try out new ways of living, should in the long run be better able to differentiate assigned roles from inner needs and values. There is a bit of empirical support for this hypothesis in a study of organizational behavior conducted by Kahn, Wolfe, Quinn, Snoek, and Rosenthal (1964). They employed the California Psychological Inventory (CPI) to measure a trait called *flexibility–rigidity*, a close cousin of Openness. Said Kahn et al., “The strong identification with his superiors and with the official goals of the organization, together with the heavy emphasis on authority relations, leads the rigid person to a rather complete internalization of his roles. His role tends to become his identity. While the flexible person may be in the course of the work day a manager, friend, personal confidant, sports fan, and explorer, the rigid person is *manager*” (p. 294; emphasis in the original). What becomes of such a person when he retires? Is he now *retired*, content to be nothing more productive? Or does he face a crisis in the loss of an identity in which he was too heavily invested? Is the departure of children from the home more stressful for the closed woman who has always regarded herself as nothing but mother? Surely these transitions must have different impacts on open and closed people.

Theorists have often noted the importance of fantasy in the acquisition of new roles. Imagining oneself as a U.S. senator, making speeches and debating foreign policy, may be important as a first step in running for election—even to a much lesser office. The housewife’s ability to see herself with an independent career probably smoothes the way for her when she begins looking for a job or returns to school. A certain flexibility in the self-concept as well as a degree of Openness would thus appear to contribute to change (Whitbourne, 1986b).

The same considerations may apply to relationships as to roles. We know that closed persons espouse traditional family ideologies (Costa & McCrae, 1978); they may also be the kinds of people who preserve a single style of relating to significant others throughout their lives. For them, perhaps, Father and Mother remain the only real adults, to be loved or feared but never recognized as people with their own limitations. For them, perhaps, wives remain homemakers and husbands breadwinners, and woe to a spouse who fails to live up to these expectations or tries to exceed them! For them, perhaps, children never grow up and are always to be advised or scolded or praised according to how well they live up to parental standards.

Roger Gould’s work on adult development is concerned with the

achievement of adult consciousness: The profound realization that we control our own lives and must take responsibility for them; that life is not necessarily fair and railing against injustice will not make it so; that we ourselves are not innocents but have a darker side; that death will come to us as to our parents. Gould sees these insights emerging over the first half of life in a regular sequence. But are they universal? Or is it perhaps only the open individual who sees and feels enough of life to make these discoveries and reach the real freedom that freedom from illusion allows?

The Life Narrative

Some personality psychologists have turned their attention to *life narratives* or *stories*—the stories people tell about their lives (Bertaux, 1981; McAdams, Josselson, & Lieblich, 2001). These researchers are generally not concerned with the historical accuracy of the accounts, but rather with their psychological significance. Life stories may reveal inner dreams or fears that help individuals make sense of their own lives. In this sense, they are akin to Thematic Apperception Test (TAT) stories: People may project their inner needs into the story of their lives.

Although some theorists have reservations about relating life stories to personality traits (McAdams, 1996), believing each is entitled to its own research program, it is obvious to us that life narrative data might fruitfully be analyzed in terms of enduring dispositions. Surely the individual's perception of his or her life as revealed in a life story will be colored by prevailing emotions and experiential styles. For example, one of Conley's (1985) respondents who scored as intuitive (or high in Openness) in 1935 assessed his life in 1980 as follows: "As long as I can do new things, see new places, and see new faces, living is fine. When I cannot do this for any reason, it will be enough and time to go." For this person, the meaning of life, and even its desired extent, were determined by lifelong Openness.

For most people, a life narrative is a story they tell themselves or a few others. But there is a whole literary genre of public life narrative, the autobiography. One of the greatest of these is *The Confessions* of Jean-Jacques Rousseau (1781/1953). As one of the most important figures in the development of the modern world, Rousseau has been widely studied by historians, philosophers, and political scientists. Figure 11 presents his NEO-PI-R profile, based on ratings by a political sci-

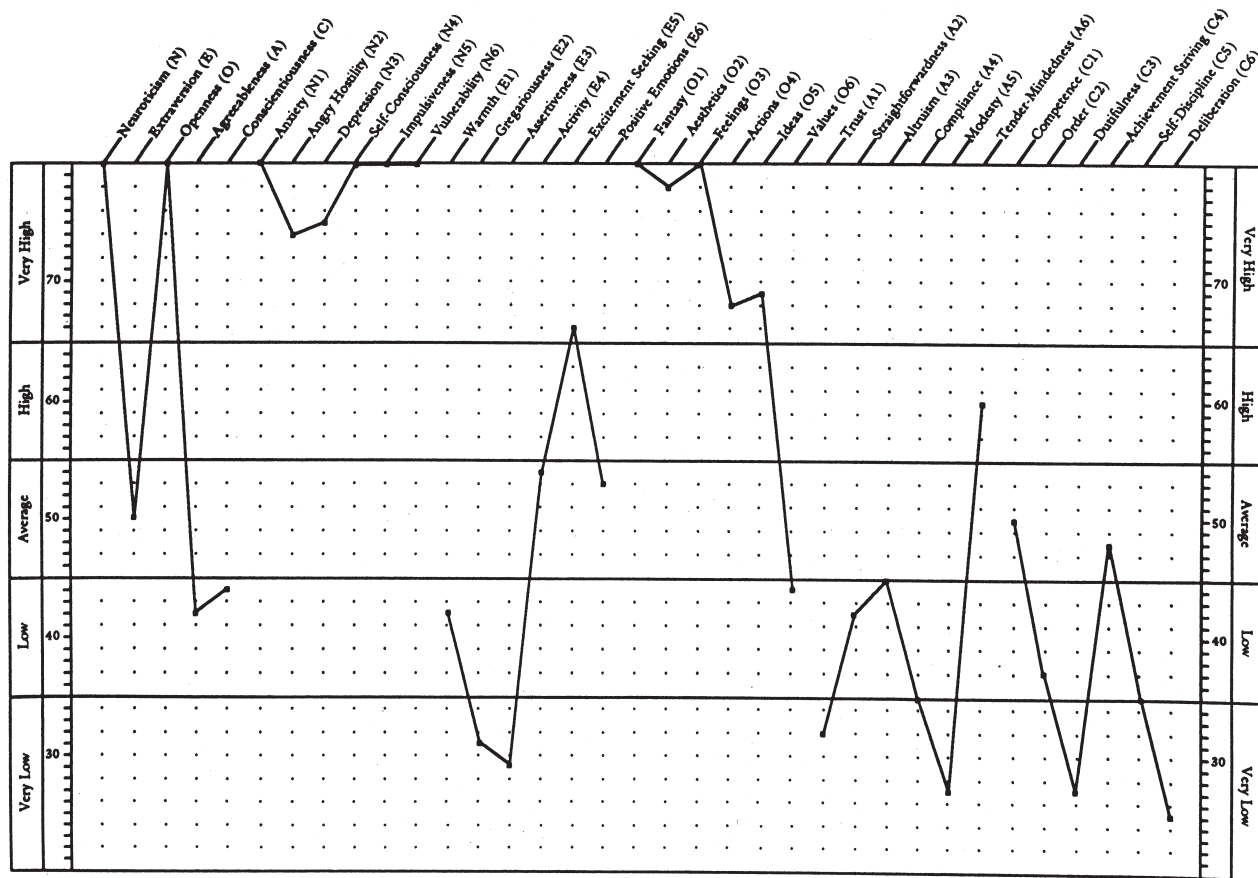


FIGURE 11. NEO-PI-R profile of Jean-Jacques Rousseau, as rated by A. M. Melzer. Profile form reproduced by special permission of the publisher, Psychological Assessment Resources, Inc., 16204 N. Florida Ave., Lutz, FL 33549, from the Revised NEO Personality Inventory, by Paul T. Costa, Jr., and Robert R. McCrae. Copyright 1978, 1985, 1992, by PAR, Inc. Further reproduction is prohibited without permission of PAR, Inc.

entist and Rousseau biographer, A. M. Melzer (McCrae, 1996). At a glance, one can see that Rousseau was a man of extremes, notable in particular for off-the-chart Neuroticism and Openness. One could certainly interpret many of the events in Rousseau's Objective Biography in terms of this profile. Is it surprising that one so high in Openness to Feelings should be one of the founders of the Romantic Movement in literature or that one so low in Trust but so high in Tender-mindedness would reject the political philosophy of the day and offer his own Utopian alternative in *The Social Contract* (Rousseau, 1762/1968)?

But Rousseau's personality traits affected not only his outward behavior but also his self-conception, his life narrative. In the opening line of *The Confessions* he tells readers that he has "resolved on an enterprise which has no precedent," namely, to give "a portrait in every way true to nature" (1781/1953, p. 17). And indeed, when it appeared, his autobiography was a shockingly frank self-portrait, revealing intimate details of his health and sexuality, and interpreting his life not as a series of public career moves but as a succession of personal struggles. Only an individual extraordinarily high in Openness could have conceived such a project or had such an intricately detailed life narrative on which to base it.

Unfortunately, Rousseau was characterized not only by high Openness but also by high Neuroticism and low Agreeableness. Historians confirm that his personal life was a disaster, but here life narrative and Objective Biography part company. From Rousseau's telling, he was the victim of pettiness and narrow-mindedness, a man who had been betrayed by friends and denied the credit he deserved. History tells us that he was paranoid and mistrustful, a man who alienated most of those (like the Scottish philosopher David Hume) who tried to help him. Rousseau's interpretation of his life, the way in which he sought to give purpose and meaning to his existence, was deeply colored by his enduring dispositions. Fortunately for us, his genius and his experiential Openness turned what might have been a whiny and self-serving tract into a great work of literature.

* * *

The last few years have been extraordinarily productive ones for the field of personality and aging. Longitudinal studies begun years ago by farsighted researchers have offered clear and consistent evidence of

what happens to personality dispositions with age. In the same period, personality psychology itself has revived from a period of skepticism and stagnation, bringing new vitality to an area that has proved itself indispensable in helping us to understand human beings and their lives. We have made conceptual advances and have come to see that much of what we were quarreling about was words instead of facts. A good deal has now been established, and a completely new direction for research on personality and aging can now be envisioned.

Conceptually, we have learned to distinguish personality traits from the Characteristic Adaptations that they help shape. Five-Factor Theory argues that traits are biologically based Basic Tendencies that are manifested in consistent styles of thought, feeling, and action.

The Five-Factor Model of personality has brought order to the competing systems of personality structure by showing that most traits can be understood in terms of the basic dimensions of Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. We know that we can measure these traits with an acceptable degree of accuracy by either self-reports or ratings from knowledgeable sources.

We have learned all the pitfalls of assuming that cross-sectional comparisons show age changes, and we have considered the evidence from longitudinal and sequential designs as well. And study after study has shown that, aside from small declines in Neuroticism, Extraversion, and Openness, there is neither growth nor decline in adult personality after age 30. A psychology whose sole purpose was to explain how personality changes with age would have little to say. Indeed, it becomes more pertinent to explain how personality remains stable.

For stable it is, not only in groups, but in the rank order of individuals. Longitudinal studies using a variety of instruments, and using raters as well as individuals' own reports, find great continuity in the level of traits in individuals. These same findings are confirmed by retrospective accounts that show the same traits active in youth and in old age.

But the stability of traits does not imply that life itself must be repetitious and stagnant. Lives change, history moves on, and all of us must work to adapt to change or actively reshape our lives. Internally, our sense of self, our fundamental identity, may change as the social roles, values, physical attributes, and personal relationships that are central to it shift.

It is in the study of these changes that the future of personality and aging lies. Ask not how life's experiences change personality; ask instead how personality shapes lives and gives order, continuity, and predictability to the life course, as well as creating or accommodating change. For the psychologist as well as the aging individual, enduring dispositions form a basis for understanding and guiding emerging lives.