3. A plea for mechanisms

JON ELSTER

Introduction

Are there lawlike generalizations in the social sciences? If not, are we thrown back on mere description and narrative? In my opinion, the answer to both questions is No. The main task of this essay is to explain and illustrate the idea of a *mechanism* as intermediate between laws and descriptions. Roughly speaking, mechanisms are *frequently occurring and easily recognizable causal patterns that are triggered under generally unknown conditions or with indeterminate consequences*. They allow us to explain but not to predict. An example from George Vaillant gives a flavor of the idea: ''Perhaps for every child who becomes alcoholic in response to an alcoholic environment, another eschews alcohol in response to the same environment'' (Vaillant 1983, p. 65). Both reactions embody mechanisms: doing what your parents do and doing the opposite of what they do. We cannot tell ahead of time what will become of the child of an alcoholic, but if he or she turns out either a teetotaler or an alcoholic, we may suspect we know why.

Although the bulk of this essay concerns the use of mechanisms in the social sciences, the idea has wider application. In her claim that ''the laws of physics lie,'' Nancy Cartwright uses the following illustration:

> Last year I planted camellias in my garden. I know that camellias like rich soil, so I planted them in composted manure. On the other hand,

I am grateful to Nancy Cartwright, G. A. Cohen, Robyn Dawes, Dagfinn Føllesal, Peter Hedström, George Loewenstein, Richard Posner, Nils Roll-Hansen, Bernt Stigum, and the late Amos Tversky for comments on an earlier version of this essay. I also benefited from comments by the participants in the conference on ''Mechanisms'' in Stockholm in June 1996, notably by my discussant Arthur Stinchcombe. A fuller version of the present essay is presented in my *Alchemies of the Mind* (forthcoming).

> the manure was still warm, and I also know camellia roots cannot take high temperatures. So I did not know what to expect. But when many of my camellias died, despite otherwise perfect care, I knew what went wrong. The camellias died because they were planted in hot soil. . . .
>
> So we have an explanation for the death of my camellias. But it is not an explanation from any true covering law. There is no law that says that camellias just like mine, planted in soil which is both hot and rich, die. To the contrary, they do not all die. Some thrive; and probably those that do, do so *because* of the richness of the soil they were planted in. We may insist that there must be some differentiating factor which brings the case under a covering law: in soil which is rich and hot, camellias of one kind die; those of another thrive. I will not deny that there may be such a covering law. I merely repeat that our ability to give this humdrum explanation precedes our knowledge of that law. On the Day of Judgment, when all laws are known, these may suffice to explain all phenomena. But in the meantime we do give explanations; and it is the job of science to tell us what kinds of explanations are admissible. (Cartwright 1983, pp. 51–2)

Cartwright's example relies on what I shall call *type B mechanisms*. Briefly defined, they arise when we can predict the triggering of two causal chains that affect an independent variable in opposite directions, leaving the net effect indeterminate. I contrast them with *type A mechanisms*, which arise when the indeterminacy concerns which (if any) of several causal chains will be triggered. An example from the natural sciences of type A mechanisms can be taken from fear-elicited behavior in animals.[1] Environmental stimuli can trigger one of three mutually incompatible fear reactions: fight, flight, or freeze. We know something about the conditions that will trigger these reactions. Thus "in response to a painful shock, animals will typically show increased activity, run, jump, scream, hiss or attack a suitable target (e.g., another animal) in their vicinity; but, in response to a stimulus associated with shock, the animal will most likely freeze and remain silent. The brain mechanisms that mediate these two kinds of reactions are quite distinct" (Gray 1991, p. 244). But although we can identify the conditions that trigger freeze versus either fight or flight, we do not know which will trigger fight versus flight. "Rather than

[1] I am indebted to Nils Roll-Hansen for suggesting this example.

thinking in terms of two systems for reaction to different classes of punishment, it makes better sense to imagine a single fight/flight mechanism which receives information about all punishments and then issues commands *either* for fight *or* for flight depending on the total stimulus context in which punishment is received'' (ibid., p. 255). But to say that the independent variable is ''the total stimulus context'' is equivalent to saying that the two responses are triggered under ''generally unknown conditions.'' Cartwright's example and the flight–fight example provide robust instances of mechanisms in the natural sciences.

In developing the idea of a mechanism, I shall proceed as follows. In the following section, I provide a more precise definition of the notion of a mechanism. In the third section, I discuss some pairs of psychological mechanisms in more detail. In the fourth section, I indicate how these elementary mechanisms may form building blocks in constructing more complex explanations. In the fifth section, I discuss some conditions under which it may be possible to move beyond the ex post identification of mechanisms to predictive statements ex ante. The final section offers a few conclusions.

Explaining by mechanisms

Let me begin by clearing up a terminological ambiguity. In *Explaining Technical Change*, I used the term ''mechanism'' in a sense that differs from the one adopted here (1983a). In that work, I advocated *the search for mechanisms* as more or less synonymous with the reductionist strategy in science. The explanation of cell biology in terms of chemistry or of chemistry in terms of physics are strikingly successful instances of the general strategy of explaining complex phenomena in terms of their individual components. In the social sciences, this search for mechanisms (or for ''microfoundations'') is closely connected with the program of *methodological individualism* – the idea that all social phenomena can be explained in terms of individuals and their behavior.

In that earlier analysis, the antonym of a mechanism is a *black box*. To invent an example at random, suppose somebody asserted that unemployment causes wars of aggression and adduced evidence for a strong correlation between the two phenomena. We would hardly accept this as a lawlike generalization that could be used in explaining specific wars unless we were provided with a glimpse inside the black box and told *how* un-

employment causes wars. Is it because unemployment induces political leaders to seek for new markets through wars? Or because they believe that unemployment creates social unrest that must be directed toward an external enemy, to prevent revolutionary movements at home? Or because they believe that the armament industry can absorb unemployment? Although many such stories are conceivable, some kind of story must be told for the explanation to be convincing, whereby "story" I mean "law-like generalization at a lower level of aggregation."

In the present analysis, the antonym of a mechanism is a scientific *law*. A law asserts that given certain initial conditions, an event of a given type (the cause) will always produce an event of some other type (the effect). For example, if we keep consumer incomes constant, an increase in the price of a good will cause less of it to be sold ("the law of demand"). Again, we may ask for a story to support the law. One story could be that consumers maximize utility. Gary Becker (1962) showed, however, that the law of demand could also be supported by other stories (e.g., that consumers follow tradition, as far as possible, or even that they behave randomly).

In more abstract terms, a law has the form "If conditions C_1, C_2, ... C_n obtain, then always E." A covering-law explanation amounts to explaining an instance of E by demonstrating the presence of C_1, C_2 ... C_n. At the same abstract level, a statement about mechanisms might be "If C_1, C_2 ... C_n obtain, then sometimes E." For explanatory purposes, this may not seem very promising. It is true, for instance, that when there is an eclipse of the moon, it sometimes rains the next day, yet we would not adduce the former fact to explain the latter. But consider the idea that when people would like a certain proposition to be true, they sometimes end up believing it to be true. In this case, we often do cite the former fact to explain the latter, relying on the familiar mechanism of wishful thinking.

This is not a lawlike phenomenon. Most people entertain some beliefs that they would like to be false. Ex ante, we cannot predict when they will engage in wishful thinking – but when they do, we can recognize it after the fact. Of course, the mere fact that people adopt a belief that they would like to be true does not show that they have fallen victim to wishful thinking. Even if the belief is false or (more relevantly) inconsistent with information available to them, we cannot infer that this mechanism is at work. To draw that conclusion, more analysis is needed. Is this a regular

pattern in their behavior? Do they often stick to their beliefs even as evidence to the contrary becomes overwhelmingly strong? Do they seem to be strongly emotionally attached to their beliefs? Can other hypotheses be discarded? By standard procedures of this kind, we can conclude, at least provisionally, that wishful thinking was indeed at work on this particular occasion. In doing so, we have offered an explanation of why people came to hold the belief in question. The mechanism provides an explanation because it is *more general* than the phenomenon that it subsumes.

In my earlier terminology, going from a black-box regularity to a mechanism is to go from "If *A*, then always *B*" to "If *A*, then always *C*, *D*, and *B*." In this perspective, mechanisms are good because their finer grain enables us to provide better explanations. Understanding the details of the causal story reduces the risk of spurious explanations (i.e., of mistaking correlation for causation). Also, knowing the fine grain is intrinsically more satisfactory for the mind. (On both points, see Elster 1983a, Ch. 1.) On the view set out here, the move from theory to mechanism is from "If *A*, then always *B*" to "If *A*, then sometimes *B*." (Because fine grain is desirable in itself, I also urge the further move to "If *A*, then sometimes *C*, *D*, and *B*.") In this perspective, mechanisms are good only because they enable us to explain when generalizations break down. They are not desirable in themselves, only *faute de mieux*. Yet because the best is so hard to attain, it can easily become the enemy of the good. The "plea for mechanisms" is not an argument against lawlike explanations, only against the idea that when such explanations fail – which they usually do – we must fall back on narrative and description.

Mechanisms often come in pairs. For instance, when people would like the world to be different from what it is, wishful thinking is not the only mechanism of adjustment. Sometimes, as in the story of the fox and the sour grapes, people adjust by changing their desires rather than their beliefs (Elster 1983b). But we cannot make a lawlike statement to the effect that "Whenever people are in a situation where rational principles of belief formation would induce a belief that they would like to be false, they either fall victim to wishful thinking or to adaptive preference formation." To repeat, most people entertain some beliefs they would like to be false. Or take another pair of mechanisms: adaptive preferences versus counter-adaptive preferences (sour grapes versus forbidden fruit). Both phenomena are well known and easily recognizable: Some people prefer what they

can have, while others tend to want what they do not or cannot have. Yet it would be absurd to assert that all people fall in one of these two categories. Similarly, some people are conformists, some are anticonformists (they always do the opposite of what others do), and some are neither.

When the paired mechanisms, as in most of the examples given so far, are mutually exclusive, they are what I called *type A mechanisms*. Yet paired mechanisms can also operate simultaneously, with opposite effects on the dependent variable. Even when the triggering of these mechanisms is predictable, their net effect may not be. These are what I call *type B mechanisms*. For an example, consider the impact of taxes on the supply of labor:

> A high marginal tax rate lowers the opportunity cost or "price" of leisure, and, as with any commodity whose price is reduced, thereby encourages people to consume more of it (and thus do less work). But, on the other hand, it also lowers peoples' incomes, and thereby may induce them to work harder so as to maintain their standard of living. These two effects – the substitution and income effects, in economists parlance – operate in opposite directions, and their net effect is impossible to predict from theory alone. (Le Grand 1982, p. 148)

As in Cartwright's camellia example, the separate effects are robust propensities, but the net effect is more contingent. The *indeterminacy* associated with mechanisms can, therefore, take two forms. With type A mechanisms, we may not be able to predict whether they will be triggered; with type B mechanisms, we may not be able to assess the net effect of two opposing mechanisms.

A further distinction may be made between cases in which the two opposing mechanisms are triggered simultaneously by the same cause, and cases in which one is triggered by the other.[2] I shall refer to these as mechanisms of type B_1 and B_2, respectively. A paradigm case of a B_2 mechanism is the "opponent-process system" (Solomon and Corbit 1974). An initial experience of pleasure or pain, when terminated, instead of bringing the subject back to the preexperience baseline state, generates an oppositely signed experience of pain or pleasure. Euphoria and withdrawal in drug addiction illustrate the pleasure–pain sequence. The pain–

[2] An application of this distinction of Marx's theory of the falling rate of profit is in Elster (1985, pp. 123–4).

pleasure sequence is illustrated by the relief a woman experiences upon learning that her fear of cancer was ungrounded.

I have asserted that we cannot tell, in general, when a given mechanism will be triggered or, in the case of several mechanisms that operate simultaneously or successively, what their net effect will be. In doing so, I may appear to dismiss a large psychological literature demonstrating the operation of these mechanisms under specific conditions. Consider, for instance, the availability and representativeness heuristics (Tversky and Kahneman 1974). For each of these mechanisms, it is possible to specify conditions under which it will predictably come into play. Yet this set of sufficient conditions, which can be realized in experimental situations, may not often appear in real-life cases. Knowing that C_1, $C_2 \ldots C_4$ are sufficient for X to occur and D_1, $D_2 \ldots D_5$ are sufficient for Y to occur does not help us to predict what will happen in the presence of C_1, C_3, D_2, D_4. If we know that "If C_1, then sometimes X" and "If D_4, then sometimes Y," we should be ready for either effect. In fact, in some conditions, both the availability and representativeness heuristics are observed:

> When in a game there is a 50% chance of winning, people expect that a small number of rounds will also reflect this even chance. This is only possible when runs of gains and losses are short: a run of six losses would upset the local representativeness. This mechanism may explain the well-known gamblers' fallacy: the expectation that the probability of winning increases with the length of an ongoing run of losses. The representativeness heuristic predicts that players will increase their bet after a run of losses, and decrease it after a run of gains. This is indeed what about half the players at blackjack tables do. . . . But the other half show the reverse behavior: they increase their bets after winning, and decrease them after losing, which is predicted by the availability heuristic. After a run of losses, losing becomes the better available outcome, which may cause an overestimation of the probability of losing. [The] repertoire of heuristics predicts both an increase and decrease of bet size after losing, and *without further indications about conditions that determine preferences for heuristics, the whole theoretical context will be destined to provide explanations on the basis of hindsight only.* (Wagenaar 1988, p. 13, italics added)

To summarize, I am not advancing explanation by mechanisms as an ideal or a norm. Explanation by laws is better – but also more difficult, usually

too difficult. Moreover, as will be clear by now, I am not suggesting that mechanisms can be identified by formal conditions analogous to those that enter into the formulation of laws. "If p, then sometimes q" is a near-useless insight. Explanation by mechanisms works when and because we can identify a particular causal pattern that we can recognize across situations and that provides an intelligible answer to the question, "Why did he do *that*?"

Some elementary mechanisms

In this section, I offer a more systematic discussion of some elementary or atomic mechanisms. The purpose of the discussion is to demonstrate the range and power of mechanism reasoning. I am not trying to prove any particular thesis, only to persuade the reader of the fruitfulness of the approach. I first consider two type A mechanisms and then two type B mechanisms.

Adaptive preferences versus wishful thinking

In Festinger's theory of cognitive dissonance (Festinger 1957, 1964; Wicklund and Brehm 1976), dissonance is stipulated to arise when a person holds two or more "cognitions" that are inconsistent with one another. Here, cognitions include not only ordinary factual beliefs but also consciously held values as well as mental representations of the choices or behaviors of the subject. The notion of inconsistency is based on "expectations about what goes with what . . . built up on the basis of past experience, including notions of logical relations, cultural mores, and learned empirical correlations among events" (Festinger and Bramel 1962, p. 255). Thus if a person has just bought a car of brand X, the expectation is that he will not believe brand Y to be better. Or to take another famous paradigm of dissonance research, if a subject is asked to write an essay giving arguments for abortion and chooses to do so even if the rewards are small, expectations are thwarted by learning that he is actually strongly against abortion.

Dissonance reduction (or avoidance) takes place by changing or blocking some of the dissonant cognitions, and sometimes by adding new ones. In spite of certain ambiguities in Festinger's original formulations, the process has to be thought of as unconscious. In the car ex-

ample, for instance, dissonance can be avoided by reading ads for the car one has just bought and avoiding reading ads for other brands. These behaviors cannot proceed from conscious choices for the purpose of reducing dissonance, for if one *knew* that the ads were read or avoided to bolster one's confidence in the choice one had just made, no bolstering could take place. Somehow, one "just gravitates" toward the behaviors that confirm the wisdom of the choice. The dissonance-reducing change in the essay-writing example is even more obviously constrained to be unconscious. When subjects are unable to tell themselves that the behavior is justified by their lack of choice or by a high reward, they reduce their dissonance by adopting a more favorable attitude toward the view they defended.

In an important special case, dissonance is generated by the presence of a desire that X be the case and a belief or suspicion that X is not the case. There are (at least) five possible outcomes. (1) People can try to modify the world to make X be the case. (2) They can accept the fact that the world is not as they want it to be. (3) The beliefs may change so that they acquire a firm belief that X is in fact the case. (4) The desires may change so that they cease to desire that X be the case. (5) The desires may be changed so that they come to desire that X not be the case ("sour grapes"). Of these, (1), (2), and (3) may represent autonomous behaviors or mental processes, governed by the reality principle rather than the pleasure principle. In particular, (4) may result from autonomous character planning such as has been advocated by Stoics, Buddhists, and others. By contrast, (3) and (5) are escape mechanisms that operate at an unconscious level.

As far as I know, nothing is known about when dissonance reduction takes the form of wishful thinking (3) and when it appears as adaptive preference formation (5). Note that each reaction, although valuable in easing short-term tension, has undesirable – and different – long-term consequences. If the wishful thinking leads to the formation of false beliefs about the world, as it usually although not necessarily does, acting on these beliefs can have bad consequences. Adaptive preferences tend to overshoot by a kind of psychic momentum that carries them beyond mere indifference into aversion (Veyne 1976, p. 312; Mora 1987, p. 72). The point I want to make is that *it matters* which of the two functionally equivalent mechanisms is triggered, because each of them has further, different consequences over and above that of reducing the tension.

Spillover, compensation, and crowding out

In an essay "How to profit from one's enemies," Plutarch observes "that a man is farthest removed from envying the good fortune of his friends or the success of his relatives, if he has acquired the habit of commending his enemies, and feeling no pang and cherishing no grudge when they prosper." This illustrates what I shall call *the spillover effect*: envy of one's enemies tends to induce envy of one's friends. A few pages later he notes that "since all human nature bears its crop of contention, jealousy and envy . . . , a man would profit in no moderate degree by venting these emotions upon his enemies, and turning the course of such discharges, so to speak, as far away from his associates and relatives." This is what I shall call *the compensation effect*: envy of one's enemies immunizes against envy of one's friends.

More formally, the spillover effect is that if a person follows a certain pattern of behavior P in one sphere of his life, X, he will also follow P in sphere Y. The compensation effect is that if he does not follow P in X, he will do so in Y. To these we should add the *crowding-out effect*: If he does follow P in X, he will not do so in Y. If the compensation effect and the crowding-out effect obtain simultaneously, they yield a *zero-sum effect*.[3]

Tocqueville's analyses of American democracy rely heavily on these mechanisms and on their interaction. Rather than reproducing what I have written elsewhere on this topic (Elster 1993, Ch. 4), I shall give some examples from other subject matters and other writers. I begin with an example from discussions of participatory democracy. First, there is the thesis advocated by Carole Pateman (1970): If people participate in decision making at the workplace, they will also become more predisposed to participate in politics. This is the spillover effect. Second, there is what we may call the Oscar Wilde thesis: Even under socialism, Wilde ob-

[3] Claims that mental life in general is subject to a zero-sum law amounts to a theory – the "hydraulic theory of the mind" – rather than to a mechanism. It is, moreover, a false theory, as acutely noted by Tocqueville: "It would seem that civilized people, when restrained from political action, should turn with that much more interest to the literary pleasures. Yet nothing of the sort happens. Literature remains as insensitive and fruitless as politics. Those who believe that by making people withdraw from greater objects they will devote more energy to those activities that are still allowed to them treat the human mind along false and mechanical laws. In a steam engine or a hydraulic machine smaller wheels will turn smoother and quicker as power to them is diverted from the larger wheels. But such mechanical rules do not apply to the human spirit" (Tocqueville 1986, p. 168).

served, the week will only have seven evenings, implying that participation in one sphere will be at the expense of participation in other spheres. This is the crowding-out effect. Third, one might argue that people have a need to participate in joint decision-making processes, so that if they are denied, say, democracy at the workplace, there will be a strong demand for political democracy, and vice versa. This is the compensation effect.

In an article on the organization of leisure, Harold Wilensky (1960) traces what he calls "the compensatory leisure hypothesis" and "the spillover leisure hypothesis" back to Engels's work, *The Conditions of the Working-Class in England in 1844*. The first states that the worker who is alienated at work compensates by active and energetic leisure activities; the second that "he develops a spillover leisure routine in which alienation from work becomes alienation from life; the mental stultification produced by his labour permeates his leisure." Rather than assuming that the one or the other mechanism is true always and everywhere, we may conjecture that some individuals are subject to the first and others to the second, or that the same individual cycles between the two. A conjunction of the two mechanisms might offer a more satisfactory account than either of them taken separately.

So far I have considered spillover and compensation as intrapersonal mechanisms of attitude formation. The last few remarks suggest, however, that one may enlarge the perspective to consider how similar effects may be at work in interpersonal relations. When young aristocrats and young elite commoners are educated together, the compensation effect may dampen the dueling tendencies of the former while the spillover effect may enhance those of the latter (Billacois 1990, p. 136). Another example is provided by individual donations to charity. A spillover-like mechanism is that embodied in the *norm* of fairness: If others give more, I should give more, too (Elster 1989, p. 187 ff.; Sugden 1984). A compensation-like mechanism arises from more outcome-oriented utilitarian reasoning: If others give more, my contribution matters less so that I can give less (Elster 1989, p. 46 ff.; Margolis 1982). I will return to this example shortly.

Contrast effect versus endowment effect

In the mid-1980s, Amos Tversky suggested (personal communication) that past experience has a dual effect on present welfare. On the one hand,

there is an endowment effect: A memory of a good experience is a good memory, the memory of a bad one a bad memory. Hence a good past tends to improve the present, a bad past to make it worse. On the other hand, there is a contrast effect: A good experience in the past tends to devalue less good experiences in the present, and a bad event in the past will similarly throw the present into favorable relief. A meal at a superlatively good French restaurant may cause one to enjoy later meals at French restaurants (and perhaps at other restaurants, too) less than one would otherwise have done. Conversely, there is nothing like recovery from illness to make you appreciate a normal state of health.[4]

Given the existence and regular operation of these two mechanisms, we can ask several questions. First, there is a mechanism of type B_2: If a positive or negative experience triggers a negative or positive contrast effect (and assuming no endowment effect), will the net effect be positive or negative? This question has been much discussed ever since it was recognized that "if the best can come only rarely, it is better not to include it in the range of experiences at all" (Parducci 1968, p. 90).[5] Second, there is a mechanism of type B_1: What is the net effect, mediated by contrast and endowment, of experiences at an earlier time on welfare at a later time? This was the question identified by Tversky. Third, combining the first two questions, we might ask about the net effect of the initial experience on welfare overall, either as discounted to the earlier time or without discounting. In a given case, the net effect on later welfare might be negative (a negative contrast effect being stronger than a positive endowment effect), but the net effect on overall welfare might still be positive (the positive utility from the experience itself offsetting the negative net effect at the later times). To my knowledge, nobody has studied the third and more important question.

In a study of the second question, Amos Tversky and Dale Griffin (1991) assume that the contrast effect, unlike the endowment effect, re-

[4] These phenomena should not be confused with the opponent-process mechanism (see section 2). In that process, an initial positive experience generates a later negative experience independently of whatever other events may transpire. In the presently discussed case, the subsequent effects depend on later events. If all my later meals are taken in superlatively good French restaurants, the contrast effect will not operate.

[5] Conversely, he argues that "The ideal lower end-point might be a strong electric shock, unbearable, but quickly over. The shock would have to be readministered occasionally, whenever it dropped from the context or whenever its memory ceased to be dreadful" (Parducci 1984, p. 16).

quires some similarity between the present and the past. The superlative French meal will not, for instance, tend to devalue a meal in a Chinese restaurant. In this specific case, that assumption seems reasonable. If, however, we imagine a man in prison dwelling miserably on how it felt to be free or a recovering patient enjoying his improved health, we do not need to stipulate a contrast between specific types of experience. Given this assumption and the high susceptibility of judgments of similarity to framing effects, they note that "one should find ways to treat the positive experiences of the past as different from the present" (ibid., p. 299). They also note, however, that people may not have much freedom in the framing of hedonic events. I will return to that issue in the fifth section, "From mechanisms to laws."

The bulk of Tversky and Griffin's study is devoted to an analysis of net effects in specific experimental situations. Although they assert that their predictions were confirmed, this turns out to mean mainly that if the past events were dissimilar from the present ones, there was no contrast effect. In addition, they note that the principle of loss aversion suggests a prediction (which was confirmed) that the negative contrast effect following a high payoff will be larger than the positive contrast effect following a low payoff (ibid., p. 305). They do not, however, offer any prior reasons for believing that the contrast effect will dominate the endowment effect or vice versa when both operate. It turns out that in one of their two experiments the endowment effect was stronger, whereas in the other, the two effects were of roughly equal strength. Although loss aversion is cited as an explanation for the difference between the two experiments, no explanation is given for the results obtained in any one of them.

Later, George Loewenstein and I (Elster and Loewenstein 1992) generalized Tversky's idea to a larger variety of experiences. In addition to endowment and contrast effects that arise from one's own past experiences, we identified similar effects that arise from the anticipation of one's future experiences, from other people's experiences, and from merely imagined or counterfactual experiences. Because the term "endowment" does not fit these other contexts, we used "consumption effect" as the more general term. To some extent, we also addressed the question of the net effect. We noted that in interpersonal comparisons there is a transition from a dominant consumption effect to a dominant contrast effect that occurs at the point of equality (Loewenstein, Thompson, and Bazerman 1989). We also noted the absence of a contrast effect when the future is

expected to be worse than the present. In other cases, however, the net effect remains indeterminate. It is an open question, for instance, whether the consumption effect of daydreaming can offset the contrast effect.[6]

Consumption and contrast effects are not the only results of interpersonal comparisons. Abraham Tesser (1991) compares the painful contrast effect (or envy) with a pleasurable "reflection effect," basking in the reflected glory of a superior individual. Because both envy and reflected glory depend on our closeness to the other person, they will wax and wane together, the net effect being in general indeterminate.[7] In one of his experiments, Tesser found that the two effects were of approximately equal magnitude, with zero net effect *as far as pleasure or pain goes*. Yet this finding does not imply that this condition is equivalent to one in which the subject and the comparison person are equal, in which case both effects would be zero. The latter condition would produce not only zero net pleasure or pain but also zero arousal. Tesser found, however, that the subjects in the former condition did experience arousal, as evidenced in their enhanced ability to perform simple tasks and decreased ability to perform complex tasks. I return to some methodological implications of this finding in the final section of the chapter.

Desires and opportunities

Actions are caused by desires and opportunities. But the explanation of behavior need not stop there. We may go one step further and inquire into the causes of the causes. In some cases, the desires are caused by the opportunities. In others, desires and opportunities have a common cause in an antecedent variable. I shall discuss both cases, with reference to the "Tocqueville effect" in the explanation of revolutionary behavior. In a dynamic version, the effect says that discontent with existing conditions increases when conditions improve. The static version is that discontent

[6] See Elster (1997, Ch. IV. 1) for some comments on daydreaming. Note that in daydreaming, the consumption effect comes first and the contrast effect later, upon return to reality. If people discount the future, therefore, they might indulge in daydreaming, therefore even if on balance it makes them worse off.

[7] According to Ben-Ze'ev (1992, p. 568), "Achievements of those very close to us evoke pride *rather than* envy when these achievements are . . . connected with us in such a manner that we can share the credits they bestow" (my italics). Thus he asserts, in my terminology, that closeness is the triggering variable in a type A mechanism. Tesser, by contrast, asserts that closeness is part of a set of conditions that induce both pride and envy in a type B mechanism. A priori, one cannot tell who is right – or whether both might sometimes be.

is greater when conditions are better. Although Tocqueville (1955, p. 176) runs the dynamic and the static effects together, they are clearly distinct; either might exist without the other.[8] I first discuss the dynamic and then the static effect.

The standard account of the dynamic Tocqueville effect is probably that when opportunities increase, aspiration levels increase even faster, making for more discontent. The idea lacks, I think, the compelling simplicity one would want to have in a mechanism. More satisfactory is Tocqueville's idea that "the mere fact that certain abuses have been remedied draws attention to the others and they now appear more galling" (1955, p. 177). Also, economic progress makes for more occasions for abuse, by bringing more individuals into contact with the inefficient state administration (ibid., pp. 178–9). Moreover, as suggested by Hirschman and Rothschild (1973, p. 46), economic progress that is not accompanied by ascent along other dimensions may create a frustrating state of status incongruence.

The possibility of telling different fine-grained stories to support the dynamic Tocqueville effect illustrates the move from "If *A*, then sometimes *B*" to "If *A*, then sometimes *C*, *D*, and *B*" (see the second section of this chapter). Whichever of the stories we prefer, it seems clear that the dynamic Tocqueville effect may but not need to go together with a net increase in discontent. After all, economic satisfaction may offset the frustration caused by dealings with state bureaucrats or by status incongruence. Tocqueville does not offer a theory to the effect that economic progress invariably causes revolution but rather an argument to the effect that it may do so. The status-incongruence version shows this especially clearly. While economic progress satisfies one desire, it creates another and leaves it unsatisfied. The net effect of an increase in opportunities on satisfaction and on the desire for further change can go either way.

Consider next the static effect – the relationship between hardship and change. I have suggested elsewhere (Elster 1985, pp. 352–3) that necessity may be not only the mother of invention but also an obstacle to invention. Although invention requires motivation, which is stimulated by necessity, it often requires resources that may be lacking in situations of hardship. A similar two-pronged argument applies to collective action, and more specifically to revolutionary behavior. Revolutions are rarely caused by extreme hardship, because people living at subsistence conditions have to

[8] See Elster (1989, p. 68) for a similar distinction in the analysis of wage bargaining.

spend all their time simply staying alive. They may have the desire for change but no opportunities to effect it. Conversely, those well-off may have the opportunities but not the desire. In between, there may be a range of incomes that have a positive net effect – mediated by desires and opportunities – on the propensity to engage in revolutionary behavior. Although the static Tocqueville effect cannot be monotonic throughout the whole income range, the tendency for middle peasants to be more revolutionary than landless peasants indicates that it may be monotonic in the lower part of the range. Even in that range, however, the sign of the net effect is in general indeterminate, although the sign of the first derivative is not.

The static and dynamic effects may obviously be combined. When people grow richer, their frustration may increase; at the same time, their increased wealth may give them the resources to do something about their dissatisfaction. I now proceed to a more general discussion of such cases.

Molecular mechanisms

In this section, I go beyond elementary or atomic mechanisms to molecular mechanisms, both at the intrapersonal and the interpersonal levels. The usefulness of the mechanism approach is, I believe, particularly apparent in the analysis of complex psychic and social phenomena. The purpose is to illustrate and stimulate the imagination rather than to argue for any specific thesis.

The idea of molecular intrapersonal mechanisms can be illustrated by the following example. Suppose that you have been with a lover for a while but that he or she decides to break off the relationship. Because of the contrast effect, there will be an initial reaction of grief. You may then observe your mind play the following trick on you: To reduce the pain of separation, you redescribe your lover to yourself so that he or she appears much less attractive. This, obviously, is a case of sour grapes, or adaptive preference formation. You then notice, however, that the endowment effect is also affected. By degrading the other, you can no longer enjoy the memory of the good times you had together. In fact, you will feel like a fool thinking back on the relationship you had with an unworthy person. To restore the good memories, you have to upvalue the other, but then, of course, the grief hits you again.

The exact course of events will depend on the relative strength of the

different mechanisms at work. Just as people "may vary in the degree to which their reactions are dominated by endowment or by contrast" (Tversky and Griffin 1991, p. 298), they may also differ in their susceptibility to adaptive preference formation. A person dominated by the contrast effect and highly vulnerable to the sour grapes mechanism will initially be very miserable and then quickly overcome the grief. A person dominated by the endowment effect will not suffer so much in the first place. Others may be miserable for a long time, and still others may experience cycles of misery and relief. And if we add counteradaptive preference formation to the range of mechanisms, even more possibilities come into play. Such interplay of mechanisms is the stuff of novels and of everyday life. Perhaps it is time for the social sciences to consider them?

Tocqueville relies heavily on molecular interpersonal mechanisms. In the *Ancien Régime*, he plays on both the compensation effect and the spillover effect in his explanation of the radical character of the French Revolution. Because of the lack of political freedom under the old regime, "the political ferment was canalized (*refoulé*) into literature, the result being that our writers now became the leaders of public opinion and played for a while the part which normally, in free countries, falls to the professional politician" (Tocqueville 1955, p. 142): This is the compensation effect. Later, "when the time came for action, these literary propensities were imported into the political arena" (ibid., p. 147): This is the spillover effect.

Another Tocquevillian example concerns the relation between religion and politics. If a society has a democratic political organization, does that make it more or less likely to be strongly religious? On the one hand, there is a *compensation effect*: "I doubt whether man can support complete religious independence and entire political liberty at the same. I am led to think that if he has no faith, he must obey, and if he is free he must believe" (Tocqueville 1969, p. 444). In other words, when people's need for authority is not satisfied in politics, they seek it in religion. On the other hand, there is a *spillover effect*. "Men who live in times of equality find it hard to place the intellectual authority to which they submit, beyond and outside humanity. . . . One can anticipate that democratic peoples will not easily believe in divine missions, that they will be quick to laugh at new prophets, and that they will wish to find the chief arbiter of their beliefs within, and not beyond, the limits of their kind" (ibid., p. 435). Here, the argument is that the lack of authority in politics tends to under-

mine religious authority rather than support it. As noted previously, there is no need to see these analyses as contradicting each other. In fact, their conjunction may provide a better explanation of the fate of religion in democratic societies than either does separately.

A common theme in *Democracy in America* is that the flaws of democracy can be overcome by more democracy; democracy secretes the antidotes to its own diseases. An important special case of this argument stipulates that democracy may reduce the desire of the citizens to do what democracy allows them to do. We have just seen that religion may be an endogenous product of democracy, through the compensation effect. Religion, in turn, will limit the desires of the citizens in a way that may counteract the greater opportunities for licentious or dangerous behavior that democracy offers them. Thus "while the law allows the American people to do everything, there are things which religion prevents them from imagining and forbids them to dare" (ibid., p. 292). The argument does not allow us, however, to conclude anything about the net effect. If the opportunity set is greatly expanded and the desires only weakly restrained, the net effect of democracy may be to increase rather than to reduce the incidence of the behavior in question. The two pairs of mechanisms are summarily represented in Figure 3.1: *If* the influence of democracy on religion is mediated by the compensation effect rather than the spillover effect, democratic societies will be religious. *If* the negative effect of democracy on desires (mediated by religion) is strong enough to offset the positive effect of democracy on opportunities, democratic citizens will behave moderately.

From mechanisms to laws

Although it is difficult to establish laws in the social sciences, that goal will always, for better or for worse, continue to guide scholars. In this section, I discuss some ways of going beyond mechanisms to lawlike statements.

Eliminating spurious mechanisms

In some cases, the presence of two opposed mechanisms may be an artifact of social perception. Consider "Like attracts like" versus "Opposites attract each other." These apparently opposed proverbs may in fact turn out

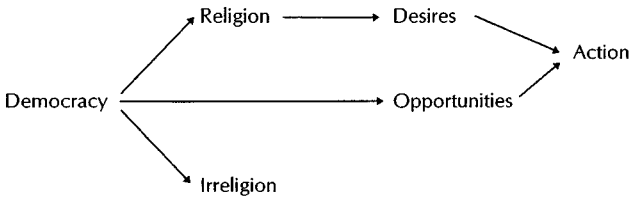

Figure 3.1. Mechanisms for interaction of democracy and religion

to be different versions of the same claim – analogous to the glass that is both half-full and half-empty. If people's curiosity and thirst for novelty is triggered by options that are neither very similar to nor very dissimilar from one's present state (Middleton 1986), their marital choices might be uniquely guided by the search for an optimal difference between their spouse and themselves (Byrne and Kurmen 1988). Depending on the perspective, that difference might be seen as closer to similarity or to dissimilarity, giving rise to the two opposed proverbs.

The point can also be put in a slightly different way. If attractiveness is an inversely U-shaped function of novelty or similarity, each of the two opposing mechanisms might simply describe different parts of the curve. "On the rising point of such a curve, increased liking is held to result from increases in the independent variable (e.g., unexpectedness, complexity). Ultimately, some optimal level is reached, whereafter increases in the independent variable are held to give rise to reductions in liking. Thus, up to a point, 'the more the merrier,' after which, 'one can never have too much of a good thing' " (Ortony, Clore, and Collins 1988, p. 166).

Predicting mechanisms from outcomes

I have been assuming that mechanisms shape outcomes, but it may also be the other way around. Consider again donations to charity. Earlier, I identified two mechanisms that can be summarized as "Give much when others give much" and "Give little when others give much." An indeterminacy then arises if we are unable to predict which individuals in which situations will be subject to the one or the other reaction. We could, however, look at the problem the other way, and assume that people (1) would like to give as little as possible but (2) would also like to tell a

story (i.e., cite a mechanism) to others and to themselves that justifies small donations. We can then predict that small donations by others will trigger the fairness mechanism and large donations, the utilitarian mechanism. The outcome is the same in both cases (viz. small donations). This identity obtains not because different motivations yield the same outcome (as in Becker's argument) but because people adopt the motivation that will yield the desired outcome.

This example is a bit awkward, because if donations are always going to be low, it is not clear that there could ever be an occasion for releasing the utilitarian mechanism. In other examples that I now proceed to cite, this difficulty does not arise. Let me first cite a Jewish joke about anti-Semitism.

Ignace Paderewski, Poland's post–World War I premier, was discussing his country's problems with President Woodrow Wilson:

> "If our demands are not met at the conference table," he said, "I can foresee serious trouble in my country. Why, my people will be so irritated that many of them will go out and massacre the Jews."
>
> "And what will happen if your demands are granted?" asked President Wilson.
>
> "Why, my people will be so happy that they will get drunk and go out and massacre the Jews." (Telushkin 1992, p. 112)

Similarly, studies of gambling have "found that, like . . . winners, losers increased the riskiness of subsequent bets" (Greenberg and Weiner 1966, reported in Cornish 1978, p. 17). If you win, you can afford to take bigger risks; if you lose, you increase the odds to recoup your losses. It is also significant that in the Twenty Questions developed by Gamblers Anonymous to help problem gamblers diagnose themselves, all the following appear:

- After losing, do you feel you must return as soon as possible and win back your losses?
- After you win, do you have a strong urge to return and win more?
- Do arguments, disappointments, or frustrations create within you an urge to gamble?
- Do you have an urge to celebrate any good fortune by a few hours of gambling?

Other addictive behaviors, such as smoking or drinking, have similar features: They are triggered by bad news or bad moods as well as by good

news and by good moods. In either case, "This calls for a drink" or "This calls for a cigarette" is cited as the justification for indulging one's craving.

Amos Tversky and Eldar Shafir (1992) conducted a series of experiments that are also relevant in this connection. One of them, which is related to gambling, finds that a majority of subjects assert that they will accept a second gamble if they won in a prior gamble *and* if they have lost in a prior gamble; however, only a minority say they will accept a second gamble if they do not know whether they will have won or lost in the first. As they observe, this is a violation of the sure-thing principle, which states that if x is preferred to y knowing that event A obtained, and if x is preferred to y knowing that A did not obtain, then x should be preferred to y even when it is not known whether A obtained. Their explanation for the observed violation of this principle is cognitive, not motivational. But at least in the gambling example – and assuming that the subjects like the thrill of gambling and do not only think in financial terms – a motivational explanation could also be possible. If one really wants to gamble but knows that it is not a good idea, one needs an excuse, a reason, a story to justify doing so. Winning will provide one story, losing will provide another, but ignorance does not. One cannot decide to accept the gamble by telling oneself that whatever happens in the first gamble, one *will have* an excuse for continuing, because that is not how excuses work. They are not planned ahead of time; rather, one observes the situation when it arises and finds a reason in it to do what one wants to do.

To the extent that mechanisms provide one with excuses for doing what one would like to do, we can predict which mechanism will in fact be triggered under which conditions. The effect is a little bit like "hedonic framing." The hypothesis of hedonic framing states that "people edit gambles in a way that would make the prospects appear most pleasant" (Thaler and Johnson 1990, p. 53). In other words, hedonic framing involves a preference-based choice among different ways of describing the same situation. Similarly, the would-be minimizer of charitable donations compares the fairness mechanism and the utilitarian mechanism and settles for the one that allows him to donate as little as possible, consistently with his need to retain his self-respect. In both cases, the comparison and choice would have to take place unconsciously: One cannot *decide* to trick oneself in these ways. A difference between the two effects can be brought out by citing an objection to hedonic framing: "Imagine you had just received an unexpected gain of $50. This could be hedonically reframed

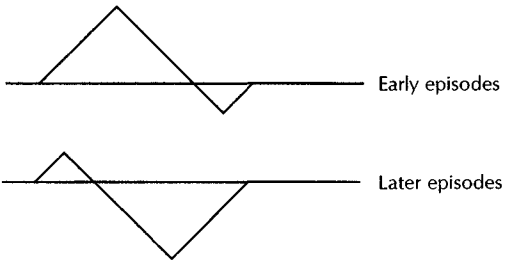

Figure 3.2. Dominance of main and opponent effects over time.

into two gains of $25, but why stop there? Why not 50 gains of $1?'' (ibid., p. 56). By contrast, the hedonic manipulation of mechanisms as excuses is limited by the small number of stories that are available.

Identifying the triggers

Consider ''Absence makes the heart grow fonder'' versus ''Out of sight, out of mind.'' La Rochefoucauld pointed to a possible triggering factor that would explain when the one or the other mechanism would be observed: ''Absence lessens moderate passions and intensifies great ones, as the wind blows out a candle but fans up a fire'' (Maxim #276). Here, the generalizing strategy is to identify a particular aspect of the *situation* that allows us to predict which mechanism will be triggered. Similarly, Tversky and Griffin show that the endowment effect dominates the contrast effect when the present event differs qualitatively from the past one. With regard to the opponent-process effect (see the second section of this chapter), Solomon argues that the main effect dominates in the initial episodes, and the opponent effect in later episodes. In addiction, for instance, euphoria initially dominates withdrawal, which then comes to dominate in later stages (see Figure 3.2).

In other cases, we might be able to point to properties of the *individual* that allow us to predict the triggering of a particular mechanism. As Tversky and Griffin also mention, some individuals may be more sensitive to the contrast effect than to the endowment effect, and perhaps we might be able to identify them on the basis of other properties. Yet if Walter Mischel (1968) is right in his claim that there is little intrapersonal, cross-situational consistency, such differences might themselves be situation-

specific. Moreover, the behavior might not be rigidly fixed to a given type of situation. If Tocqueville is right in arguing that people have a need for a sphere in which they are independent and also for a sphere in which they are subject to authority, which sphere serves which need may be a somewhat arbitrary matter.

A more ambitious strategy for anchoring mechanisms in laws relies on catastrophe theory. When I gave a talk on this topic some years ago, I cited various pairs of opposed mechanisms – sour grapes versus forbidden fruits, the attraction of likes versus the attraction of opposites, conformism versus anticonformism. Normal Schofield then remarked in the discussion that this kind of *bifurcation* is exactly what one would expect in cusp catastrophe models. In these models, the surface describing the behavior of a dependent variable as a function of two independent variables folds in on itself in a cusp. Within a certain range, a given constellation of the independent variables is thus consistent with several values of the dependent variable. Moreover, these values tend to be far apart from each other, corresponding to the polarized nature of mechanisms.

More recently, Abraham Tesser and John Achee (1994) have developed this argument more systematically. They observe that in many social situations, the function relating the independent variables to the dependent variable is two-valued rather than one-valued; hence the distribution of behaviors is bimodal rather than unimodal. Jack Brehm's (1966) theory of "reactance," for instance, is based on the premise that social pressure can decrease as well as increase conformity. John Roemer's (1985) idea of the "psychology of tyranny" is also relevant here. The tyrant induces fear in his subjects but also hatred. The former makes them less likely to rebel, the latter more likely. Tesser and Achee argue, however, that the indeterminacy disappears once we go beyond state variables and introduce path dependence or hysteresis:

> Dissonance theory provides a very nice psychological model for hysteresis. Assume that one's disposition is consonant with engaging in the behavior and that undergoing negative social pressure is dissonant with engaging in the behavior; one's disposition is dissonant with not engaging in the behavior, and the presence of negative social pressure is consonant with not engaging in the behavior. If one starts out high on the behavior in the face of strong social pressure, then as one's disposition decreases, dissonance increases. To reduce the dissonance, one

> will look for additional cognitions to support the behavior. Hence the behavior will tend to remain high even in the face of a decreasing disposition. On the other hand, starting with strong social pressure and low levels of behavior, increasing one's disposition will increase dissonance. To reduce the dissonance, one will look for additional cognitions to support not engaging in the behavior. Hence the behavior will remain low even though the disposition is increasing. (Tesser and Achee 1994, p. 104)

The model has several further implications. In the first case, as the disposition continues to decrease in the face of strong social pressure, there will come a point when the person switches from engaging in the behavior to not engaging in it. In the second case, as the disposition continues to increase, there will come a point when the person switches from not engaging in the behavior to engaging in it. Moreover, the level of disposition at which the first switch occurs is lower than the level at which the second occurs. A person who has adopted an unpopular opinion will need to see a lot of the evidence for it fritter away before he gives it up, whereas an uncommitted person will need a lot of evidence for it before adopting it. Finally, a given combination of social pressure and disposition can lead to high as well as low engagement in the behavior, depending on where the person initially started up.

Many of the arguments offered by Tesser and Achee are tantalizingly similar to the ideas I have been developing here. It may indeed turn out to be the case that pairs of opposed mechanisms correspond to different parts of the cusp surface. In that case, we could use knowledge of the past behavior of the individual to go beyond mechanisms and predict what he will do. This would still fall short of the ideal of science, which is to predict and explain using state variables only. Appealing to past values of the variables in order to explain behavior in the present is intrinsically unsatisfactory (Elster 1976). Although we would prefer to explain in terms of the traces left by the past in the present rather than in terms of the past itself, this approach would at least provide a determinate explanation.

Yet ultimately I think the two approaches are quite different. Consider the following discussion of two opposed reactions to social pressure:

> J. W. Brehm suggested the presence of a motive to maintain one's freedom to behave as one wishes. This countermotive to conformity is termed *reactance*. There is now a substantial body of literature docu-

> menting the operation of this motive. In one study, for example, Heilman gave subjects on the streets of New York the opportunity to sign a petition for an issue they mildly endorsed. In the course of the interaction, some of the subjects learned that someone else believed that people should not be allowed to sign such petitions. This latter group was more likely to sign the petitions than were subjects who were not exposed to this social pressure. So, sometimes social pressure encourages contrary behavior. (Tesser and Achee 1994, pp. 103–4)

In this approach, the contrary of conformism is behaving as one wishes, or nonconformism. In my thinking about mechanisms, the contrary of conformism is anticonformism – doing the opposite of what others do or try to get you to do. Elsewhere I have described these antonyms of conformism as its external and internal negation respectively (Elster 1993, Ch. 2). The person who stands up to pressure, and disregards what others think if he believes he is right, is autonomous. As La Bruyère observed, however, "there is equal weakness in opposing the mode and in embracing it" (*The Characters* XIII.11). The person who always does the opposite of what others do or want him to do is as heteronomous – dependent on others – as the conformist is (Elster 1983b, pp. 23, 67). In the catastrophe model, the opposite of adaptive preference formation would presumably be the absence of any causal influence of the feasible set on the preferences. In my approach, the antonym is counteradaptive preference formation. Although the catastrophe model may be capable of explaining when we do or don't bend to pressure, it does not seem capable of explaining why we sometimes bend over in the opposite direction.

A plea for disaggregation

When opposing explanation by mechanisms to explanation by laws, I have assumed that the latter is invariably deterministic. Much social science, however, relies on statistical explanation, a procedure notoriously plagued by many conceptual difficulties. One cannot use statistical explanation to account for individual cases, although it is often used in that way. Also, in this mode of analysis, it is particularly difficult to distinguish causation from correlation. I believe the mechanism approach provides yet another reason why statistical explanations tend to be weak and unreliable.

Suppose that a scholar decides to study the dependence of donations to

charity on the amount of money donated (and known to be donated) by other people, and that there turns out to be very little correlation. It might be tempting to conclude that people do not really take account of how much others give when deciding how much to give themselves. An alternative explanation might be that the population consists of two roughly equal-sized groups, one motivated by the norm of fairness and one motivated by more utilitarian considerations. On this account, *everybody* would look to others before deciding how much to give but would differ in the way the decisions of others affect their own. To uncover the presence of these two opposed mechanisms (nonlawlike tendencies), one has to go to a lower level of aggregation and look inside the black box.

This perspective suggests a reinterpretation of the Mischel's findings. Contrary to what would be implied by a universal spillover effect, people who are altruistic, aggressive, or impulsive in one context (e.g., work) do not systematically behave the same way in other settings (e.g., the family). It does not follow, however, that there is no causal relationship operating across contexts. It might be the case that what we observe is the net effect of spillover and compensation. Suppose, for instance, that we found a relatively weak correlation between individual rates of time discounting across different activities or for different goods. The explanation might be that for some individuals the habit of foresight spills over from one sphere to other spheres, whereas for others the demands of self-control are so strenuous that when they achieve it in one part of their life they have to give themselves a break elsewhere.

Similarly, it has often been observed that human beings are subject to two very strong desires: the desire to be like others and the desire to differ from others, conformism and anticonformism. If some individuals are strongly dominated by the former desire and others by the latter, the aggregate effect might be very weak, suggesting that people are mostly autonomous rather than heteronomous. Theories of voting behavior, for instance, have identified both an underdog mechanism and a bandwagon mechanism (Simon 1954). Those subject to the former tend to vote for the candidate who is behind in preelection polls, whereas those subject to the latter vote for the front-runner. With many voting for the underdog, the frontrunner might lose, and vice versa. If the two types are more evenly mixed, there might be no noticeable net effect, so that the polls would be good predictors of the actual vote. The lack of influence of polls on voting

in the aggregate does not show, however, that individuals are unaffected by the polls. The neutral aggregate could mask a homogeneous population of neutral individuals – or a heterogeneous population of individuals who are all strongly affected but in opposite directions.

George Vaillant (1983, p. 65) observes that in the aggregate, "there is no evidence that [various mediating factors] statistically increase the risk of alcohol abuse in children if they are not biologically related to the alcoholic family member." Yet, as he goes on to say in the statement cited in the opening paragraph of this chapter, this weak aggregate effect could mask two strong, oppositely directed effects at a less aggregate level. If that is in fact the case, strategies of intervention might be justified that would be pointless if children were never or rarely driven to alcoholism because their parents drink. This is perhaps the most important implication of the argument. For research purposes, the disaggregate approach may be too expensive or otherwise impractical. For purposes of public policy, however, identification of subgroups may be crucial.

The plea for disaggregation also has consequences for the interpersonal case. In the third section of this chapter, "Some elementary mechanisms," I discussed Tesser's findings that the conjunction of the contrast effect and the reflection effect may yield an emotional state that is neutral as far as pleasure and pain goes. To predict behavior, however, we may need to know the strength of each mechanism, not only their net effect. Type B mechanisms within individuals may neutralize each other, as may type A mechanisms across individuals, but that does not allow us to infer that they are absent. Nor can we assume that the net effect is all that matters for prediction or intervention.

References

Becker, G. 1962. "Irrational behavior and economic theory," *Journal of Political Economy* 70, 1–13.

Ben-Ze'ev, A. 1992. "Envy and inequality," *Journal of Philosophy* LXXXIX, pp. 551–81.

Billacois, F. 1990. *The duel*. New Haven: Yale University Press.

Brehm, J. 1966. *A theory of psychological reactance*. New York: Academic Press.

Byrne, D., and K. Kurmen. 1988. "Maintaining loving relationships," in R. J. Sternberg and M. L. Barnes (eds.), *The psychology of love*. New Haven: Yale University Press.

Cartwright, N. 1983. *How the laws of physics lie*. Oxford University Press.
Cornish, D. B. 1978. *Gambling: A review of the literature and its implications for policy and research*. London: Her Majesty's Stationery Office.
Elster, J. 1976. "A note on hysteresis in the social sciences," *Synthese* 33, 371–91.
Elster, J. 1983a. *Explaining technical change*. Cambridge University Press.
Elster, J. 1983b. *Sour grapes*. Cambridge University Press.
Elster, J. 1985. *Making sense of Marx*. Cambridge University Press.
Elster, J. 1989. *The cement of society*. Cambridge University Press.
Elster, J. 1993. *Political psychology*. Cambridge University Press.
Elster, J., and G. Loewenstein. 1992. "Utility from memory and anticipation," in G. Loewenstein and J. Elster (eds.), *Choice over time*. New York: Russell Sage.
Festinger, L. 1957. *A theory of cognitive dissonance*. Stanford University Press.
Festinger, L. 1964. *Conflict, decision and dissonance*. Stanford University Press.
Festinger, L., and D. Bramel. 1962. "The reactions of humans to cognitive dissonance," in A. J. Bachrach (ed.), *Experimental foundations of clinical psychology*. New York: Basic Books.
Gray, J. A. 1991. *The psychology of fear and stress*. Cambridge University Press.
Greenberg, M. E., and B. B. Weiner. 1966. "Effects of reinforcement history upon risk-taking behavior," *Journal of Experimental Psychology* 71, 587–92.
Hirschman, A., and M. Rothschild. 1973. "The changing tolerance for income inequality in the course of economic development," *Quarterly Journal of Economics* 87, 544–65.
Le Grand, J. 1982. *The strategy of equality*, London: Allen and Unwin.
Loewenstein, G., L. Thompson, and M. Bazerman. 1989. "Decision making in interpersonal contexts," *Journal of Personality and Social Psychology* 57, 426–41.
Margolis, H. 1982. *Selfishness, altruism and rationality*. Cambridge University Press.
Middleton, E. 1986. "Some testable implications of a preference for subjective novelty," *Kyklos* 39, 397–418.
Mischel, W. 1968. *Personality and assessment*. New York: Wiley.
Mora, G. F. de la. 1987. *Egalitarian envy*. New York: Paragon House.
Ortony, A., G. L. Clore, and A. Collins. 1988. *The cognitive structure of the emotions*. Cambridge University Press.
Parducci, A. 1968. "The relativism of absolute judgments," *Scientific American* (December), 84–90.
Parducci, A. 1984. "Value judgments," in J. R. Eiser (ed.), *Attitudinal judgment*. New York: Springer.
Pateman, C. 1970. *Participation and democratic theory*. Cambridge University Press.
Roemer, J. 1985. "Rationalizing revolutionary ideology," *Econometrica* 53, 85–108.

Simon, H. 1954. "Bandwagon and underdog effects in election predictions," *Quarterly Journal of Economics* 69, 99–118.

Solomon, R., and J. Corbit. 1974. "An opponent-process theory of motivation," *Psychological Review* 81: 119–45.

Sugden, R. 1984. "Reciprocity: The supply of public goods through voluntary contribution," *Economic Journal* 94, 772–87.

Telushkin, J. 1992. *Jewish humor*. New York: Morrow.

Tesser, A. 1991. "Emotion in social comparison processes," in J. Suls and T. A. Wills (eds.), *Social comparison*. Hillsdale, N.J.: Lawrence Erlbaum.

Tesser A., and J. Achee. 1994. "Aggression, love, conformity, and other social psychological catastrophes," in R. R. Vallacher and A. Nowak (eds.), *Dynamical systems in social psychology*. New York: Academic Press.

Thaler, R., and E. Johnson. 1990. "Gambling with the house money and trying to break even," *Management Science* 36, 643–660, cited after the reprint in R. Thaler, *Quasi-Rational economics*. New York: Russell Sage.

Tocqueville, A. de. 1955. *The Old Regime and the French Revolution*. New York: Anchor Books

Tocqueville, A. de. 1969. *Democracy in America*. New York: Anchor Books.

Tocqueville, A. de. 1986. *"The European revolution" and correspondence with Gobineau*. Gloucester, Mass.: Peter Smith.

Tversky, A., and G. Griffin. 1991. "Endowment and contrast in judgments of well-being," in R. Zeckhauser (ed.), *Strategy and choice*. Cambridge, Mass.: M.I.T. Press.

Tversky, A., and D. Kahneman. 1974. "Judgment under uncertainty," *Science* 185, 1124–30.

Tversky, A. and E. Shafir. 1992. "The disjunction effect in choice under uncertainty," *Psychological Science* 3, 305–9.

Vaillant, G. 1983. *The natural history of alcoholism*. Cambridge, Mass.: Harvard University Press.

Veyne, P. 1976. *Le pain et le cirque*. Paris: Seuil.

Wagenaar, W. A. 1988. *Paradoxes of gambling behaviour*. Hillsdale, N.J.: Lawrence Erlbaum.

Wicklund, R. A., and J. Brehm. 1976. *Perspectives on cognitive dissonance*. Hillsdale, N.J.: Lawrence Erlbaum.

Wilensky, H. 1960. "Work, careers, and social integration," *International Social Science Journal* 12, 543–60.