

What Might Have Been: The Social Psychology of Counterfactual Thinking

Edited by

Neal J. Roese
Northwestern University

and

James M. Olson
University of Western Ontario

Dysfunctional Implications of Counterfactual Thinking: When Alternatives to Reality Fail Us

Steven J. Sherman
Allen R. McConnell
Indiana University

Living in the here and now and focusing only on the tasks and events of the present are things that people do not do easily. Our minds wander to the past with floodings of nostalgia as we recall events, experiences, and relationships that used to be. And especially for people in the Western world, thoughts and images race to future times with feelings of hope, anticipation, or dread. This inability to stay grounded in the present or to "stop and smell the roses" was, in fact, a major theme of the counterculture of the late 1960s and early 1970s.

Related to this inability of people to keep their thoughts and feelings tied to current experiences, tasks, and issues is their inability to accept their present reality and to be satisfied with it. Instead, they have a compelling propensity to alter reality and to reflect on "what could have been," "what might have been," and "what should have been." People undo reality and find ways in which it might have been different by mutating or slipping conditions that were antecedents to the current situation. This tendency is probably no better captured than it is by Marlon Brando's lines "I could have been a contender; I could have been somebody" from *On the Waterfront*. In short, people evaluate many life events not simply by the reality of what comes to pass but also by thoughts of what might have been. This obsession with alternatives to reality, this ubiquitous imagination of other possible worlds, has been called *counterfactual thinking*, and it has become a growing area of speculation and research in psychology since the publication of a seminal paper by Kahneman and Miller in 1986.



LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
1995 Mahwah, New Jersey

In this research and theorizing about counterfactuals investigators have focused on several interesting and important issues. For example, in one major line of research, investigators focused on discovering which aspects of reality are most changeable or mutable (e.g., Gleicher et al., 1990; Miller & Gunasegaram, 1990). A second area of interest has involved the important affective implications of counterfactual generation (e.g., Miller, Turnbull, & McFarland, 1990; Niedenthal, Tangney, & Gavanski, 1994). Finally, there has been a good deal of work relating counterfactual thought to perceptions of causality both through specific causal constructions (Wells & Gavanski, 1989) and in more general formulations of attribution theory (Lipe, 1991).

Most recently, researchers in counterfactual thought have started to focus on the psychological functions served by a person's generating alternatives to reality. The previous chapter in this volume is, in fact, devoted to a detailed discussion of the functions of counterfactual thinking (Roese & Olson, chapter 6). The assumption is that counterfactuals must serve some very important psychological functions if they are so prevalent. The identification and understanding of these functions can help psychologists to explain why people are so compelled to imagine alternative realities.

Most theorists agree that there are three major functions served by counterfactual generation. The first of these is affect regulation. The imagination of other possible worlds can help people feel better about current reality and life's circumstances. This function is generally served by the generation of *downward counterfactuals*, that is, alternatives that are worse than what exists. Thus, after a car accident, the consequences of which were a broken leg and \$5,000 damage, one might think, "If my car had swerved one more foot to the left, I would have been killed by the oncoming truck." Generating worse-world scenarios such as this makes current reality seem not so bad. That negative-affect regulation is an important function of counterfactual thought is indicated by the results of work by several researchers, who have reported a tendency for study participants to feel better following the generation of downward counterfactuals to negative experiences (Markman, Gavanski, Sherman, & McMullen, 1993; Roese, 1994).

The second important function served by counterfactual thought is preparation for the future. When life is not as good as people like or when they have not met their goals, they may focus on how to avoid these same negative outcomes or failures in the future. By generating counterfactual alternatives to reality, people can think about how outcomes could have been different, and by extension, how they might be different in the future. This preparation function is best served by generating *upward counterfactuals*, alternatives that would have led to better

outcomes. Although upward counterfactuals might leave one feeling dissatisfied and blameworthy (because the negative outcome could have been avoided), they allow one to think of different courses of action, courses that should lead to happier outcomes in the future. Such counterfactual generations may be especially worthwhile when similar decisions and circumstances are likely to present themselves in the future (Markman et al., 1993) or when the outcome was perceived as controllable and, thus, changeable (Roese & Olson, 1995). Indeed, Roese (1994) found that a person's generating upward counterfactuals following failure in an anagram task led to intentions to perform success-facilitating behaviors and to actual behavioral improvement in a subsequent task.

Related to preparation for the future, a third function of counterfactual generation has recently been considered, the induction of feelings of controllability (Markman, Gavanski, Sherman, & McMullen, 1995; McMullen, Markman, & Gavanski, chapter 5). Counterfactual generation is intimately tied to perceptions of causality (Wells & Gavanski, 1989). And once a causal representation of an event is developed, the event sequence seems more sensible, predictable, and controllable. Perceiving the world as controllable and predictable is perhaps *the* major motivation in human judgment and behavior (Langer, 1975). This need can be so strong that illusions of control develop for situations in which there is clearly no objective control. In fact, the need to feel in control of one's circumstances and outcomes may lead to such illusions of control even when these perceptions carry with them terrible affective consequences.

The following is an anecdotal case. A father in Detroit had planned for his young daughter to fly to Phoenix to visit her grandparents. At the airport, the girl cried and said she didn't want to go; she was afraid to fly and to be alone on the plane. The father explained that she would be well taken care of and that the grandparents would meet her on the other end. She would have a wonderful time. The plane crashed, and the girl was killed. It was difficult for the father to believe that a random, uncontrollable event could cause his daughter's death. He blamed himself and believed that he had control over what had happened: "If only I had realized how dangerous this really was. . . ." At last report, the father was still severely depressed and unable to function normally. Similarly and more generally, rape victims and victims of serious diseases often blame themselves for what happened to them (Bulman & Wortman, 1977; Janoff-Bulman, 1979; S. Taylor, Lichtman, & Wood, 1984; Wood, Saltzberg, Neale, Stone, & Rachmiel, 1990). Despite the terrible guilt and self-directed anger that arises from such attributions, the belief that events are not arbitrary and that the actors could have done things that would have changed the outcomes ("if only . . .") is apparently worth all the grief.

This need for control implies that the events within a scenario that one had control over will be more mutable than events that were not under one's control. Markman et al. (1995) found exactly this pattern—counterfactuals to outcomes in a gambling situation were most often generated for those aspects of the setting that involved one's own decisions. Also consistent with the idea that counterfactual generation serves the function of imparting feelings of control are results reported by Roese and Olson (1993b). They found that participants used *subtractive* counterfactual changes (mutations that remove a factual feature in reconstructing reality) when they undid successful outcomes but used *additive* mutations (mutations that add new antecedents) to undo failures. These kinds of mutations are consistent with the person's need to perceive control. Such mutations imply that doing things is what leads to success and that the only reason for failure is that one did not exert control and act in a necessary way.

It is clear, then, that counterfactual thinking can and does serve positive psychological functions. Counterfactuals can repair bad feelings or prevent the onset of such feelings, they can give one hope and confidence in preparing for the future and even play a role in improving that future, and counterfactuals can allow one to feel a sense of control over life events and give one a sense that the world and its outcomes are knowable and predictable.

Interestingly, a focus on the functions of counterfactual thinking seems to have left psychologists with the feeling that counterfactuals may be the greatest thing since ESPN. They are the panacea to brighten a person's day, to give people hope, to empower them, and to improve their circumstances. Even when the downside of counterfactual generation is recognized, recent evidence seems to indicate that this downside is minimally disturbing and disruptive. For example, upward counterfactuals can lead to negative affect as one realizes that a current negative reality need not have occurred and could have been averted. Yet, findings indicate that these negative feelings may not necessarily accompany upward counterfactuals (Roese, 1994) and that such negative feelings are especially unlikely for individuals with high self-esteem (Roese & Olson, 1993a) or for people who are chronically focused on the future or who are induced to think about the future (Boninger, Gleicher, & Strathman, 1994). Thus, counterfactuals can allow people to have their cake and eat it, too—to be better prepared for the future, to feel more empowered and in control, and to feel better about their current circumstances (or at least not to feel worse).

In the remainder of this chapter, we consider the possibility that there may be a darker side to counterfactual generation, a dysfunctional side that has potentially serious negative effects on emotional states, behavior, and judgments. In general, counterfactual thinking represents a way of

problem solving and a way of organizing and understanding experiences. Thus, counterfactual generation falls in the same broad category as the use of heuristic principles, the development of schemas or scripts, and the adoption of categorization techniques. The benefits of using heuristics, developing schemas, and forming categories or stereotypes are well known. They help people to simplify a complex world and to improve the efficiency and effectiveness of thought and action. But the costs of these processes and principles are also well known. They leave people subject to biases and to errors of judgment that can manifest themselves in unfortunate and self-defeating behaviors and perceptions. In the same way, psychologists should recognize that counterfactual thinking can serve some very positive functions but that it also leaves people open to thoughts and judgments that can be costly emotionally and practically. We attempt to outline and analyze some of these dysfunctions, without losing sight of the positive functions that counterfactual generation can serve.

ENDING UP WITH THE WRONG CAUSAL INFERENCE

Discussions of the functionality of counterfactuals have made clear that an important component of this functionality is the generation of a counterfactual that helps one understand the causal structure of a chain of events in an action sequence so that a changed outcome in the future is a possibility (e.g., Markman et al., 1993, 1995; Roese, 1994). Unless the causal inference based on the counterfactual generation is correct, however, this functional value of counterfactuals is questionable. And, unfortunately, the ability of people to arrive at an accurate causal analysis of an event sequence is far from perfect. The research literature is replete with examples of people displaying systematic biases in their interpretations of events and explanations of causal relations. Further, the growing evidence of mutation rules (i.e., factors that influence the mutability of events in a causal sequence) has indicated that these biases in causal analysis may inhere in the counterfactuals that such rules produce. People may mutate event features that played no role in the actual outcome and may fail to mutate features that were critical for the outcome to occur. Because of the role of counterfactual generation in causal analysis (Wells & Gavanski, 1989), the mutation of events that are not actually causal can be dysfunctional. Such counterfactuals will lead to an incorrect causal analysis and an improper understanding of the situation and may thus instigate continued poor performance in the future as well as negative affect inappropriately directed at others or even at oneself.

The results of research over the past few years have revealed a number of mutation rules that suggest when counterfactual generation is most likely to occur and which counterfactuals are the most likely to be generated. Most of these findings have been derivations and extensions of Kahneman and Miller's (1986) norm theory. In general, these rules suggest that individuals will mutate the most accessible features of a situation and not mutate features that are less accessible. More important, these mutation rules do not ensure that the features that are mutated are the actual causes of the outcome. Just as the accessibility of a feature is not an accurate indicator of that feature's frequency or probability (Tversky & Kahneman, 1973), neither is it a good indicator of causality. Yet, the counterfactual mutation of a feature implicates that feature as a primary cause of the outcome and indicates that any change in outcome in the future will best be achieved by a change in that feature.

One such bias in mutation involves the temporal sequence of events. Miller and Gunasegaram (1990) found that events that occur later in a temporal chain are more mutable than earlier events in the same sequence (but this is not true for events in a causal chain, in which the earliest items are the most mutable; Wells, B. Taylor, & Turtle, 1987). Thus, these later events are seen as more causal, and any hope of changing outcomes in the future will require that these later events be changed.

In Miller and Gunasegaram's (1990) first experiment, participants imagined themselves in the role of students in a class performing poorly on a test. They evaluated a teacher's questions critically (i.e., as being unfair) if they imagined having studied *before* the teacher wrote the questions rather than *after* the teacher wrote the questions. In other words, the teacher's poor exam (and not their own study habits) was the perceived cause of failure, due to the mutability of the late-occurring events. On the other hand, when the test was constructed prior to their studying, participants mutated their study habits and saw those as the cause of poor performance. Of course, neither analysis is rational because temporal sequence here is unrelated to cause. Yet, the counterfactual generation would have implications for future test preparation.

The temporal-sequence mutation rule can lead to unwarranted blame on the person who engages in the last action in a temporal chain. A classic example of this phenomenon is the basketball player who, with one second left on the clock, misses the winning basket. This player will be seen as the person who lost the game even if other players on the same team missed more shots and easier shots earlier in the game. Again, the ease of mutating the last event in the temporal chain in some situations leads to placing the blame on someone when others might be responsible for the ultimate outcome as much, or even more. The generation of counterfactuals also carries with it behavioral implications, such as influencing who will start in the next game.

Demonstrating another factor that affects mutation in a way that leads to an incorrect or biased causal analysis, Miller, Turnbull, and McFarland (1989) found that outcomes with an equal a priori probability of occurrence are judged to be more suspicious if there are fewer ways for the same outcome to occur. In one experiment, Miller et al. (1989) found that participants were more suspicious of the child who selected the 1 coveted chocolate-chip cookie in a jar with 19 less-desired oatmeal cookies than of the child who selected 1 chocolate-chip cookie from a jar containing 10 chocolate-chip cookies and 190 oatmeal cookies. In terms of statistical probability, selecting a chocolate-chip cookie from either jar is equally likely (5%). However, in the case of only one chocolate-chip cookie contained in the whole jar, all other alternatives result in selecting an oatmeal cookie; thus, selecting the one chocolate-chip cookie seems less likely to have occurred by chance. Consequently, the degree of suspicion (and subsequent mistrust of the child) is affected in an unwarranted way by the ease of counterfactual generation.

In addition to groundless scolding of children, Miller et al. (1990) suggest that this bias can help build and maintain stereotypes. For instance, consider two groups (a majority group of 500 members and a minority group of 50 members) that each claim to have the same small proportion of hostile members (e.g., they claim that only 2% of the members of the group are hostile). On encountering the first member of each group, you find both of them to be hostile. It is easy to imagine running into 1 of the 10 presumed hostile members of the majority group, and there is no reason to question the claim that the vast majority of this group are friendly. It is not easy to imagine running into the one and only presumed hostile member of the minority group by chance. Just as with the cookie jar scenario, fewer alternatives to the outcome make an event with fewer ways to occur (though probabilistically equivalent) seem more suspicious. Thus, a person might question the claim that only one member of the minority group is hostile and conclude that minority-group members must be more generally hostile. In this way, negative stereotypes of minority groups are more likely to be developed and maintained.

A third mutation principle involves the controllability of events. Those events over which one had control are most easily mutated and are thus most likely to be seen as causal. Of course, controllability and causality are not the same thing. Thus, people may blame themselves excessively for negative outcomes if they had control over some of the event antecedents—even if those events were in no way causal of the outcome. In a recent study involving a computer-simulated wheel-of-fortune game, Markman et al. (1995) found that participants generated more counterfactuals about the elements of the game over which they had control than about factors over which they had no control. More important, the type

of counterfactual generated involved thoughts of how a change in an aspect of the game that they had controlled could have changed the outcome of the game. These counterfactuals, in turn, would have implications for decisions that participants would make in subsequent games. The irony in this study is that participants actually had no control over the situation (i.e., the outcome was rigged). In the same way, a rape victim may mutate what she wore or where she went for a walk as ways of undoing the rape, perceiving her own controllable actions as the cause.

A fourth mutation rule is that actions, relative to nonactions, produce more counterfactuals. Again, mutations lead to inferences about causal structure (actions perceived as more causal than inactions) and thus have implications for affective consequences and subsequent behavior. Although actions are more mutable than inactions, they are in no way necessarily more causal or more important to an outcome. Thus, once again, counterfactual generations and biases in mutability can lead to an inaccurate causal understanding of a situation.

Kahneman and Tversky (1982) offered the first, now classic demonstration of the greater mutability of actions as opposed to inactions. In their scenario, two people lose the same amount of money on stock investments in Company Z. Person A considered switching to another stock but decided to keep stock in Company Z, whereas Person B had stock in another company but a year before had switched to Company Z. Kahneman and Tversky found that participants expected Person B (the individual who acted) to feel worse than Person A (the person who failed to act) about losing the money. This effect was also observed by Gleicher et al. (1990). The dysfunctional implication of this phenomenon is that one who fails to act may not seek out alternatives that could have improved a negative outcome. Failing to act can be as consequential as acting in a less-than-optimal fashion, although apparently acting will evoke more useful responses (i.e., searching for alternatives to an undesired outcome) than not acting. On the other hand, psychologists should also bear in mind the implications for positive outcomes as well. An individual may see an action that occurs before a positive outcome as causal, even if the action was unrelated to the outcome. Thus, one might develop superstitious or ritualistic behaviors in order to ensure (erroneously) future successes.

A fifth mutation rule is that people tend to mutate nonnormal situations in the direction of normal occurrences. Kahneman and Miller (1986) suggested that exceptional events are more mutable than common events. In a test of this notion, Miller, B. Taylor, and Buck (1991) found that study participants typically explained deviations in gender roles by mutating the nonnormative gender's behavior to make them seem more similar to the normative gender's behavior. In a series of experiments, Miller et al.

(1991) asked participants to explain "gender gaps" for various roles (e.g., how men and women differed with respect to voting patterns). In order to establish different default gender assumptions, they conducted pretesting, providing situations in which the normative gender was male or female, respectively. Pretest results showed that participants described the typical voter as male but the typical elementary school teacher as female. In one experiment, participants were asked to explain differences in female and male voting patterns. These participants tended to explain the differences in terms of how female voters' preferences deviated from male voters' preferences. Participants also predicted that if this gender gap in voting-pattern differences were to disappear in the future, such a change would be because female voters would act more like male voters rather than vice versa.

In another experiment involving attitudes toward elementary school teachers, participants showed the same bias for the opposite gender. They explained gender differences in terms of how men acted differently from women and suggested that an elimination of the gender gap would occur if male elementary teachers acted more like female elementary teachers rather than vice versa. According to Miller et al. (1991), perception and judgment of change are based on the differential ease of generating counterfactuals (see also Hilton & Slugoski, 1986, for similar reasoning). It is easier for one to think of nonprototypic members changing their behaviors. The actions of the prototypic group are taken as givens. The implications for subsequent decisions are striking, if not alarming. For example, the nonnormative gender would be more likely than the normative gender to be blamed for any negative occurrence related to a gender gap. Of course, any attempt to "fix" the situational outcome of counterfactual generation (see also McGill & Klein, chapter 12).

It is important to note that people sometimes do mutate mundane, normative events into exceptional ones as well. Once again, however, this differential mutability is not based on a realistic assessment of the role of an event in causing an outcome. Gavanski and Wells (1989) asked participants to consider a woman who was described either as an excellent student or as a poor student. On a particular exam, she performed either well or poorly. Thus, a good student's performing well or a poor student's performing poorly is normative, whereas a poor student's performing well or a good student's performing poorly is nonnormative. Participants were provided with both normative and nonnormative events that could be mutated to alter the outcomes. With results in contrast to the original tenets of norm theory and the findings of Miller et al. (1991), Gavanski and Wells found that participants who were provided with the normative scenarios (e.g., good student who passed the exam) mutated their alternatives *away*

from normality. Only when the outcome of the scenario was nonnormative did participants mutate the exceptional events into normal events.

Whether counterfactual mutations are toward or away from normality, in their counterfactual generations and thus in their assessments of causality, people apparently tend to use a simple heuristic principle, representativeness. People believe that there must be a similarity between any effect and its underlying cause. Thus, unusual events have exceptional causes; normal outcomes have run-of-the-mill causes. Only big causes can have big outcomes. Nisbett and Ross (1980) discussed people's tendency to assume similarity between events and their causes. The important point is that this simplifying principle, which determines the ease of counterfactual mutations, can lead to a misunderstanding of the causal structure of a situation and to incorrect implications for changing an outcome.

In some respects, many of the shortcomings of these mutation rules are reminiscent of the problems associated with biased hypothesis testing. When testing a hypothesis, people often rely on a biased subset of information in that they consider only the focal hypothesis and not alternative possibilities (Klayman & Ha, 1987; Skov & Sherman, 1986; Wason, 1968). Likewise, participants in the Miller et al. (1991) gender-gaps study might not have considered an alternative possibility—perhaps something about the normative group made them different from the nonnormative group.

In addition to qualities inherent in events themselves that render them differentially salient (e.g., temporal placement), certain characteristics of the perceiver can lead to individual differences in feature accessibility and, thus, to differences in event mutability. For example, individuals with chronically accessible constructs (Higgins, King, & Mavin, 1982) should be systematically attentive to particular stimulus features and will therefore devote greater attentional resources to these dimensions, a tendency increasing the likelihood that these features will be encoded and will guide subsequent retrieval (Kahneman & Miller, 1986; Medin & Schaffer, 1978; Nosofsky, 1987; Smith & Zarate, 1992). With such effect, chronically accessible constructs should lead to mutations along these focal stimulus dimensions because they will be especially accessible to the perceiver when generating counterfactuals.

Consider a person who demonstrates chronic accessibility for the construct gender. That person will devote great attention toward encoding events and outcomes in terms of gender. That person will therefore have a predisposition to evoke gender-related counterfactuals in order to change existing situations. For instance, such a person, considering the consequences to United States Attorney General Janet Reno after the 1993 siege of the Branch Davidian compound in Waco, Texas, may form the counterfactual, "Would Janet Reno have received such sharp criticism if

she were a man?" Consequently, counterfactuals can serve as a mechanism for the maintenance and reinforcement of chronic constructs because it should be easy for a person with a chronic expectancy to form *chronic mutations* about these well-encoded features of situations. There is one potential drawback of such a mechanism: Chronic mutations may lead a person down a biased counterfactual path, evoking many "what ifs" for features that would not have actually changed the outcome of the situation. Further, the attention devoted to a narrow range of situation features means that other dimensions that contributed more to the causal chain of events may go unnoticed, especially in situations in which one's cognitive resources are especially scant (Bargh & Thein, 1985; Bargh & Tota, 1988). Thus, perceivers may miss important causal factors in the chain of events. In the Reno case, the perceiver may focus so myopically on Reno's gender that he or she does not consider alternatives that might have led to the public criticism, for example, "Would Janet Reno have received such sharp criticism if the Davidians had not had innocent children inside the compound at the time the federal officials ended the standoff?"

In sum, it is clear that people will mutate event features of outcomes according to simple heuristic principles. When relying on these heuristics, they may often mutate events that played no causal role in the outcome or may fail to mutate events that were essential for the outcome to occur. Features of the events and even individual differences can contribute to this bias in generating counterfactuals. When perceivers thus arrive at faulty causal understandings based on their mutations, the counterfactuals they generate will have been dysfunctional.

Given that perceivers will frequently mutate noncausal events based on these systematic biases and will thus arrive at a misperception of the causal structure of a situation, an important question arises: How can perceivers avoid these pitfalls? A variety of strategies for eliminating biases in faulty judgments have been offered, and these strategies can be applied to counterfactual thinking as well. Lord, Lepper, and Preston (1984), for instance, found that asking participants to consider the opposite possibility was an effective strategy in the participants' avoiding biased processing of the value of a scientific study. In a similar vein, Kruglanski and Freund's (1983) theory of lay epistemology suggests that a person's entertaining a variety of alternatives helps him or her avoid the limiting aspects of cognitive processing. In many ways, lay epistemology and counterfactual generation share much in common in that *freezing* (the cessation of generating additional alternatives) leads to a person's accepting a state of affairs as irrevocable. Generating a counterfactual is dysfunctional when only highly accessible event features are mutated and a search for alternatives is thus prematurely terminated. For instance, when

fans decide that the basketball player who missed the final-second shot lost the game, they fail to search for additional alternatives to this highly accessible event.

Thus, it is important for individuals to generate multiple counterfactuals, rather than being blinded by the first counterfactual that comes to mind. When people mutate the most accessible antecedent event, they may be driven more by biased processing than by sound causal inference. Also, it is important for perceivers to consider not only how mutations could have changed the outcome but also how alternative antecedent events might have led to the *same* outcome. Only then will perceivers be able to judge correctly the inevitability or the avoidability of an event and to grasp the true causal structure of the situation.

UNNECESSARY NEGATIVE AFFECT

Despite the fact that some recent findings indicate that counterfactual thinking (upward counterfactuals in particular) may not always leave people feeling bad (Boninger et al., 1994; Roese, 1994), a rather substantial amount of literature documents certain negative affective consequences of counterfactual thought. Regret is a common outcome of counterfactual generation (Landman, 1993), and it is the mutations of negative outcomes in an upward direction that generally lead to feelings of regret (Landman, 1987). Similarly, dissatisfaction with one's current situation is an often-reported consequence of upward-counterfactual generation (Markman et al., 1993). Even outcomes that are positive by any absolute standard can lead to dissatisfaction if a counterfactual world in which things are even better is entertained. For example, Medvec, Madey, and Gilovich (in press) reported that silver medal winners in Olympic competition often feel dissatisfied, presumably because of the highly accessible alternative world of winning the gold medal. Interestingly, silver medal winners are often less satisfied than bronze medal winners, whose likely counterfactual world is one with no medal at all. Thus, those who objectively achieve more end up more dissatisfied, hardly a functional outcome of counterfactual thought.

In addition to inducing feelings of regret and dissatisfaction, counterfactual generation has been implicated in the arousal of shame and guilt (Niedenthal et al., 1994). When people undo a distressing outcome by changing their personal qualities (e.g., "If only I weren't so irresponsible, I wouldn't have lost my job"), they experience shame. When they undo outcomes by mutating their own specific actions (e.g., "If only I hadn't gotten drunk that night, I wouldn't have said those nasty things"), they feel guilt. Although shame and guilt may be functional emotions in motivating beneficial changes in behavior, counterfactual generations

such as those described for the father in Detroit can lead to unwarranted and debilitating feelings of guilt and shame.

Related to feelings of shame and guilt is self-blame, of which counterfactual generation is often a critical aspect. As we discussed previously, people have a strong tendency to mutate those aspects of an event sequence over which they had control. Thus, they are likely to view their own behaviors as causal of a negative outcome and to experience unwarranted blame. When one's own choices and behaviors that preceded an undesirable outcome are especially mutable, self-blame will be strong. For example, Turnbull and Mawhinney (1986, cited in Miller et al., 1990) asked participants to imagine a parking lot where they had a choice of either several parking spaces or one particular space because the rest of the lot was full. Later during the scenario, the participant's car and another car are involved in an accident when both cars pull out simultaneously and run into each other. Participants reported feeling more self-blame and said they would pay more of the other person's deductible when they had alternative parking spaces available to them than when they were forced to take the last available space in the lot. It was the ease of mutating one's own parking-space choice in the lot-not-full scenario that led these participants to view their own behavior as causal, to feel a great amount of self-blame, and to incur greater costs voluntarily.

This tendency to counterfactualize one's own behaviors that preceded negative outcomes may be especially strong for individuals with low self-esteem. Roese and Olson (1993a) asked participants to imagine that they were working with another person in various achievement domains (e.g., working with a fellow student on a joint class project, tutoring a boy in math). The outcome of the joint interaction was described as either successful or unsuccessful. After reading the scenario, participants were asked to generate counterfactuals that would change the outcome. Roese and Olson found that, following a success, participants with high self-esteem were more likely than those with low self-esteem to mutate their own actions. But following a failure, participants with low self-esteem were more likely than those with high self-esteem to mutate their own actions. These counterfactuals would of course be related to the participants' perceived locus of cause for the outcomes. Those with high self-esteem would see themselves as responsible for their successes. Those with low self-esteem would take responsibility for their failures. The nonfunctional implications of these findings are apparent. Individuals with low self-esteem may blame themselves unduly for failures, and people with high self-esteem may ignore their causal role in undesirable outcomes and not incorporate needed changes in their future behaviors. Thus, people with high self-esteem will continue to make the same mistakes, and those with low self-esteem will continually blame themselves for failures.

When the outcome with which one is trying to cope is traumatic enough or when the obsession with generating counterfactual alternatives that would have averted the event is strong enough, the self-blame that follows from these counterfactual mutations may be especially devastating. Consider again the poor father in Detroit, whose life has been destroyed not so much by the terrible tragedy involving the death of his daughter but by his inability to escape the counterfactual world in which she is still alive. Thus, counterfactuals can leave people with a double dose of bad feelings—negative affect due to the objective loss itself and further feelings of blame and regret because these losses are perceived as not having had to occur.

As we mentioned earlier, research results have shown that other kinds of traumatic events (e.g., rape, diagnosis of cancer) can lead one to ruminate and to attempt to undo outcomes that are beyond one's control (Janoff-Bulman, 1979; S. Taylor et al., 1984), with the attendant feelings of guilt and self-blame. In two field studies, Davis, Lehman, Wortman, Silver, and Thompson (1995) examined the psychological distress, ruminations, counterfactuals, and attributions of self-blame associated with adults who had lost loved ones in a variety of tragic occurrences (e.g., losing a child or spouse in a motor vehicle accident). Despite the fact that these adults played no causal role in any of the deaths, Davis et al. found that victims attempted to mutate aspects of the situation that involved their own behaviors, including trivial aspects of these behaviors that clearly had no implications for the deaths. Even when statistically controlling the analyses for general thinking about the tragedies (e.g., thinking about the accident scene), Davis et al. found a correlation between the amount of counterfactual thinking and victim distress, as measured on a series of scales designed for the assessment of affect and depression.

Davis, Lehman, Silver, Wortman, and Ellard (1994) recently examined the self-directed attributions made by spinal cord injury patients. In their study, Davis et al. asked patients who had experienced spinal cord injuries in a variety of situations (e.g., motor vehicle accidents, sporting accidents) to respond to several questions: How foreseeable was the risk of injury? Do you think about ways the accident could have been avoided (general counterfactual thinking)? Do you think about how *you* could have avoided the accident (self-implicating counterfactual thinking)? How much are you to blame for the accident (personal blame)? Davis et al. found that patients' self-implicating counterfactual thoughts were directly related to their judgments of personal blame for their injuries, independently of their global causal attributions (i.e., general counterfactual thinking) and their estimates of accident foreseeability. Thus, self-blame seems to be triggered by the production of self-implicating counterfactuals even when other factors, such as foreseeability and general causal attributions, are controlled for.

7. DYSFUNCTIONAL IMPLICATIONS

The dysfunctional implications for the self of these kinds of counterfactual generations are sobering. Searching for alternative worlds following traumatic events often results in people's mutating aspects of their own behavior that would never have changed the tragic outcomes; this tendency results in heavy self-blame. Even if changing their behavior would have changed the event outcome (e.g., a rape victim's taking a *different* street home), it is clearly irrational for one to take blame for behaviors that in foresight would not have reduced the probability of the event's occurrence. It is likely that these kinds of counterfactuals do help victims in the process of coping with tragedy and in giving them a feeling of control over their lives (Janoff-Bulman, 1979; S. Taylor et al., 1984), but the despair and depression that can result from this kind of counterfactual thinking can be devastating.

If the focus on counterfactual alternatives to a current undesirable reality becomes overly obsessive, the result can be depression. When people repetitively blame themselves for failing to achieve important goals or for physical problems or personal losses, they experience a spiraling sense of loss of personal control, resulting in feeling helpless and unable to change their current circumstances (Wood et al., 1990). Martin and Tesser (1989) discussed the phenomenon of ruminative thinking as a state in which people, after facing repeated failures in achieving their goals, focus on their failure rather than on ways to overcome their obstacles. When ruminative thinkers get bogged down in their own failures, they become unable to assess the situation in order to find helpful alternatives to their actions. Although Martin and Tesser did not discuss the role of counterfactuals, *per se*, in their theory, it seems likely that people engaged in ruminative thought would be especially likely to ruminate about the counterfactual worlds that might have been. Such a focus on thoughts of "if only . . ." and "why me . . ." can only lead to self-pity and depression and an inability to emerge from these kinds of thoughts. Counterfactual generation of this type can lead to a cycle of unproductive thinking that feeds on itself.

We have thus discussed a variety of circumstances in which counterfactual generation of one's own behaviors can lead to negative affective experiences in the form of regret, dissatisfaction, self-blame, guilt, shame, or depression. Similarly, the ease of counterfactualizing another person's behaviors that were antecedent to a negative outcome for that person can influence the amount of affect people feel for that person or the extent to which people hold the other person blameworthy. In one case, Miller and McFarland (1986) described a victim who had been severely injured during a robbery. The robbery took place either in a store that the victim usually frequented or in a store in which he rarely shopped but decided to enter on this occasion. The victim in the unusual-store scenario was

accorded more sympathy and was awarded greater compensation for his injuries than the victim in the usual-store scenario. This enhanced affect was due to the ease with which participants could generate an injury-free counterfactual world for the victim in the "unusual" store ("If only he had shopped at his regular store"). However, not all upward counterfactuals lead to enhanced sympathy for a recipient of a negative outcome. Johnson (1986) reported that participants reacted more negatively toward a woman who chose to move out of a seat and by so doing lost a large lottery that went to her original seat number than did participants who read a similar scenario in which a seat near her original position was selected in the lottery. According to Johnson, because it was easy to construct a counterfactual world in which the "seat-moving" victim had a better outcome, she was perceived as more blameworthy.

Thus, upward counterfactuals lead to enhanced affective reactions to victims of negative events. In addition, the specific affect can be quite different, depending on the circumstances surrounding the negative outcome. Two possibilities exist for explaining which type of enhanced affective reaction to victims will follow from the generation of upward counterfactuals. First, Miller et al. (1990) suggested that perceivers will try to compensate a victim in order to restore a sense of justice, but when this option is not available, they will attempt to reduce the sense of injustice by derogating the victim. In either case, compensation or derogation, the reaction will be greater when it is easy to generate a counterfactual world where the victimization did not occur. In the Miller and McFarland (1986) study, participants had the opportunity to compensate the victim. As a result, they awarded larger monetary amounts to the shopper at the unusual store, where an upward counterfactual was readily available. However, participants in the Johnson (1986) study did not have a compensation option. Thus, they attempted to restore justice by derogating the victim, more so when an upward counterfactual was easily generated. A second explanation involves the role that the upward counterfactual plays in the attribution of culpability to the victim.¹ In the Miller and McFarland (1986) study, the upward counterfactual does not causally link the actor to the negative outcome. Because the victim's behavior is perceived as unrelated to the assault, greater sympathy is offered via the upward counterfactual in the unusual-store case because the bad outcome, having been avoidable, seems much worse. However, in the Johnson (1986) study, a causal link is possible between the actor's behavior (i.e., leaving the seat) and the negative outcome. Because perceivers make a causal link between the actor's antecedent behavior and her less-than-desirable outcome, greater blame results from this upward counterfactual.

¹We thank Neal Roese for offering this explanation.

Thus, just as with the self-attributions, emotional reactions to others and behaviors toward them can be biased based on the ease of counterfactual generation and the attributional links captured in the counterfactual. The foreseeability of the outcomes or the factual bases for the decision, factors that should enter into an assessment of the quality of the decision making, may be less important than the counterfactual world that is generated. Because of this tendency, people fail to assess the rational or just levels of blame, sympathy, or compensation that a person deserves. Compensating someone with extra money just because he or she shopped at a different store is hardly an equitable decision.

Interestingly, people often react to the circumstances of others not only by counterfactualizing their behavior in the situation but by counterfactualizing their own behavior ("If we had been them"). People may blame someone who rushes to a meeting and has an accident or who follows a boss's orders to cut corners and causes the injury of a worker because they think, "If I were in that situation, I would have acted differently." Such a counterfactual comparison would be useful and functional only if the self-prediction were correct. In fact, such self-predictions are very inaccurate (Sherman, 1980). More important, they are inaccurate in a socially desirable direction. That is, people predict that they would behave far more socially desirably than they actually do. Thinking thus, people hold others to standards that they incorrectly believe they would live up to. Moreover, it is the generation of these counterfactuals about the self that leads to biases in the judgments of others. For example, people wrongly believe that they would never have shocked others if they had participated in the Milgram experiment (Sherman, 1980). (Results of analyses indicate that 67% of these participants in fact would have shocked others.) This faulty counterfactual generation leads people to perceive the typical Milgram participant as sadistic, and people would no doubt blame them and punish them for any injury caused. Once again, the easy (but perhaps inaccurate) generation of a counterfactual world can affect judgments and behaviors in a dysfunctional way.

Reacting to General Versus Specific Cases

Why can't people accept negative outcomes and circumstances as part of life and simply move on from there, thus avoiding negative affective consequences? Although psychologists have no definitive answers to this question, they can make some observations that strike us as interesting and that might point the way to answers.

We are intrigued by the fact that people seem to have no trouble at all accepting the general concept of failure or accepting a global and abstract failure rate. For example, a baseball manager would gladly accept

a preseason offer of a team average of one error per three games. Prior to a season, a basketball coach would happily accept a turnover rate of seven per game. Yet, even when the specific circumstances of any game or season are well within these acceptable limits, these same managers and coaches will react quite negatively to specific errors and turnovers and will express emotion and punish players based on these errors. The players themselves will experience regret, dissatisfaction, and guilt for errors, turnovers, and missed free throws, despite their understanding that, at a general level, some number of such failures are inevitable. In other words, people may realize and accept the fact that into every life a little rain must fall, but they will still react negatively and unacceptingly to any specific storm.

This ability to rationally accept negative occurrences at a general level and yet to experience great affect and emotion about any specific negative outcome is part of a broader tendency of people to react more emotionally to specific instances than to abstract and general cases. Thus, citizens in a community will accept very well the general principle of mixing commercial centers or low-income housing with neighborhoods—just not in their backyards. People may donate nothing at all to the starving or suffering children of the world, and yet millions of dollars will be sent to a specific child who is found abandoned or who has had the misfortune of falling down a well. In the realm of interpersonal judgment, Sears (1983) demonstrated the *person-positivity bias*, a tendency to evaluate specific individuals more favorably than the general group composed of these individuals.

There are many possible reasons that people respond to specific instances more emotionally and effectually than they do to general cases. Nisbett and Ross (1980) discussed differences in responses to concrete as opposed to abstract instances. Concrete, specific events are more memorable and attention-grabbing and thus evoke more affect than do general, hypothetical events. Tulving (1972), distinguishing between episodic and semantic representations, made a similar argument. Episodic memories (likely to be involved with specific cases) are more experiential, self-referenced, contextual, and affective. In addition, specific instances may be more likely to be processed in a central manner, whereas general cases may be processed in a more peripheral way (Petty & Cacioppo, 1984). Central processing leaves deeper and more unchangeable judgments, a consequence indicating again that more affect and emotion are associated with specific instances.

Most important for purposes of this chapter, we suggest that specific instances afford postcomputed counterfactual alternatives to be used as comparison standards. Unusual or unexpected or salient features can be mutated to yield alternative possible worlds. On the other hand, general

cases have no particular features that can be mutated. In this case, it is more likely that precomputed, preevent expectancies will be used as comparison standards (Kahneman & Miller, 1986). Therefore, the general case of seven turnovers per game affords no alternatives but to compare it to an expected average. Any specific turnover, on the other hand, affords many opportunities for change. As we have shown, the generation of counterfactual alternatives to reality is intimately involved in the level of emotional reaction.

In any case, an understanding of why specific instances (especially negative instances) spontaneously evoke "if onlys . . ." that lead to cringing, fist-pounding, and sadness can help psychologists to better comprehend the dysfunctional side of counterfactual generation. Moreover, perhaps those individuals who can accept some of the specific negative occurrences of life without spontaneous counterfactual generation and who do not lose sight of the general, larger picture are the best kinds of judges and decision makers. Those decision makers who can avoid the counterfactual alternatives to specific cases may be in the best position, emotionally and cognitively, to render good subsequent judgments. For example, most coaches and managers scream, grimace, and even throw chairs when errors or bad outcomes occur. A few coaches, on the other hand, seem to maintain their equanimity and composure under most circumstances. Tony LaRussa, the manager of the Oakland A's, does not cringe or scream when a relief pitcher is hammered or a defensive replacement makes an error. It would be interesting to correlate a manager or coach's reactivity to negative events, or better yet, to correlate their tendency to generate counterfactuals to their success rates.

COSTLY CHANGES IN BEHAVIORS

In addition to the emotional costs of counterfactual generation, there are important costs in terms of faulty subsequent decision making that is based on the counterfactuals. Consider again one of the major functions of counterfactual thinking—preparation for the future. Upward counterfactuals indicate how some change in a prior behavior or judgment might have led to a more successful outcome. By implication, if this change in behavior or judgment is in fact instituted in the future, better outcomes will occur. Furthermore, when an unsatisfactory outcome is experienced and one is likely to be in a similar decision-making situation again, upward counterfactuals are generated, and the aspects of the situation that are mutated are usually those that are under the control of the actor. Thus, changes in behaviors in the direction indicated by the counterfactuals can and will be made the next time.

This is all well and good, and in fact Roese (1994) showed that the generation of upward counterfactuals following unsatisfying outcomes can lead to improvements in future performance. However, such improvement and a better preparation for the future can be expected *only* if it is true that the unsatisfying outcome need not have occurred, *only* when the original judgment was faulty and could have been improved, and *only* when the actor can learn from previous mistakes. The problem is that bad judgments do not always lead to bad outcomes, and bad outcomes do not always indicate bad judgments. Unfortunately, counterfactuals seem to be generated in an almost spontaneous way whenever something negative occurs—regardless of the reasonableness of the prior behaviors or judgments. The manager who leaves the starting pitcher in for the last inning and loses the game is second-guessed: "If only he had brought in the relief pitcher, the game would have been won." The same manager who does bring in the relief pitcher and loses the game is again faced with a counterfactual world, this time a world where the starter stays in for the last inning and victory replaces defeat.

It is fine to recognize, after a bad outcome, that a different decision might have led to a different outcome in this specific case. It is quite another thing to assume that a different decision *should have been* made. Unfortunately, people assume that what need not have occurred ought not to have occurred, and what might have been is what should have been. Miller and Turnbull (1990) referred to this confusion of what might have been with what ought to be as the *counterfactual fallacy*. This fallacy, they argued, is based on the perception that highly mutable outcomes should not have occurred. It is the ability to generate counterfactual alternatives to an undesirable reality that creates the feeling that the negative outcome was avoidable. People seem to ignore the fact that mutability is quite different from what should have been done. Miller and Turnbull (1990) commented on how this tendency to view negative events that are easily counterfactualized as events that ought not to have happened can lead to mistaken judgments and decisions that violate the principles of justice and deservingsness.

The counterfactual fallacy can affect both judgments of others and self-relevant judgments. In fact, many of the negative affective responses directed toward the self that were discussed earlier can be understood as consequences of the counterfactual fallacy. In addition to such affective effects, this fallacy can also lead to counterproductive behaviors. Imagine playing the following game. A card is to be drawn from a normal deck. You may choose either spades, clubs, and hearts on the one hand, or diamonds on the other. If you choose correctly, you win \$100. If you choose incorrectly, you lose \$100. Wisely you choose spades, clubs, and hearts, but a diamond is drawn from the deck. You lose \$100. It is true

that if you had chosen a diamond in this specific case you would have won \$100. Does this mean that, given the choice again, you prepare better for the future by choosing a diamond or by refusing to play the game at all? Undoing the past need not be linked to changes in what we do in the future. But, as we saw earlier (Miller & Turnbull, 1990), the two often go hand in hand. A strategy, however poor, that would have led to an easily imagined success in the past will often be chosen over a better strategy that led to failure in the past. Thus, counterfactuals that would have changed a bad current reality to a better one are often assumed to be the best way to ensure a more satisfying future.

Such thinking and such a tendency to adopt upward counterfactuals in a prescriptive rather than in a descriptive way are tied to the fact that people are outcome-driven beings. People often judge the quality of decision making solely on the basis of the outcome. Therefore, strategies that are followed by success should be maintained. Strategies that are followed by failure should be changed, and the change should be in the direction of the most accessible upward counterfactual. In fact, we agree with Miller and Turnbull (1990) that it is likely to be the salience of the spontaneous upward counterfactual that causes the feeling that the judgment was bad and ought to be changed. This tendency to continue with strategies that are followed by success and to alter strategies that are followed by failure is, of course, exemplified by the win-then-stay—lose-then-switch strategies that people often exhibit during learning or hypothesis testing (Levine, 1959). This tendency can interfere with the ability to learn or solve problems quickly and efficiently.

Baron and Hershey (1988) demonstrated the strength and ubiquitousness of this outcome bias in decision making. In virtually every case that they tested, their participants judged the quality of thinking and the competence of the decision maker higher when the decision (the very same decision under the same antecedent conditions) was followed by a success rather than by a failure. In other words, participants confused fortune with wisdom, and misfortune with guilt. Similarly, Roese and Olson (1994) asked participants what options a decision maker in a World War I scenario should have chosen. Although participants understood that their judgments should not have been affected by the outcome, and they believed that their judgments were not so affected, their judgments were in fact very much outcome driven. That is, participants judged that decision makers should have known to make the decisions that, in hindsight, led to the successful outcome. Thus, sports managers will lose their jobs, CEOs will be fired, and prison furlough systems will be discontinued because of short-term or isolated, unsatisfactory outcomes. Furthermore, the presence of counterfactual alternatives to these unsatisfactory outcomes likely serve to motivate and to justify the decision to change. Although Baron and

Hershey (1988) did not collect counterfactual-generation data from their participants, we suspect that the ease and extent of counterfactual thinking is very much related to the extent of outcome bias.

The findings of Baron and Hershey are, of course, related to previous results by Walster (1966), who reported that participants attributed greater responsibility to an actor for the very same action when that action had more serious consequences. The same negligent act is judged more harshly when the consequences of that act are more severe. This relationship between outcome and culpability is a very important one in the United States system of jurisprudence. In many respects, this system of laws actually maintains that for many crimes culpability should be independent of consequence. In fact, framers of the Model Penal Code (American Law Institute, 1985) adopted this position in 1962, stating that culpability should ideally be based only on what an actor knew before an act or at the time of the act (although culpability is in fact often linked to outcome, as in murder cases). The same act committed under the same conditions should be punished equally regardless of the severity of the consequences. Furthermore, certain parts of the American judicial system follow this principle. Thus, recently in Bloomington, Indiana, a driver of a car that ran a red light and killed two pedestrians was punished by a fine of \$65. This is the standard fine for running a red light, and the consequences of the act are immaterial, provided that there was no prior intent to kill the pedestrians. Of course the public screamed in outrage and disbelief. They wanted the punishment to fit the outcome rather than the decision to run the red light, and counterfactual generations were no doubt a major part of their arguments and their affect.

The point is that negative outcomes are not simply accepted. Rarely are such outcomes viewed as unfortunate consequences of good decisions. The counterfactual, better alternative worlds are taken not only as what could have been but also as what should have been. In this light, it is clear that upward counterfactuals can better prepare people for the future only when the initial judgment was a poor one. Given people's tendency to focus on the outcomes of decisions, however, they will sometimes change good decision rules to bad ones on the basis of the counterfactuals that they generate. In fact, changing judgment strategies based on upward-counterfactual generation ought to be most beneficial for poor decision makers, but most dysfunctional for good decision makers. It may make people feel hopeful that they can ensure success in the future simply by changing to a strategy that would have had a positive outcome for a specific instance in the past, but this thinking is far from rational decision making. Good decision making requires an analysis of the conditions under which a decision was made rather than a focus on the outcome of that decision and on the counterfactual worlds that could have emerged under different judgments for the specific instance at hand.

This discussion of the role of counterfactual generation in changing behaviors or decision-making strategies is related to three other concepts that have proved to be important in the area of human judgment and behavior: hindsight bias, foreseeability, and self-handicapping.

Hindsight Bias

Hindsight bias refers to a person's judging a prior outcome as having been more predictable and even inevitable subsequent to learning the outcome (Fischhoff, 1975; Hawkins & Hastie, 1990). In other words, in retrospect, people "knew it all along." Thus, even when a priori judgments of likelihood are wrong and when decisions based on these mistaken judgments are made, people will not learn from their errors because, after the fact, they believe that they knew it in foresight and have nothing to learn.

At first blush, it seems that hindsight bias is incompatible with counterfactual generation. If outcomes seem inevitable, then counterfactual worlds are not easily conceived or generated. However, in a recent paper, Roese and Olson (1994) pointed out that hindsight bias is not only logically compatible with counterfactual generation but also that the two may even be complementary. They discussed hindsight bias as a belief that, once an outcome is known, the causal structure of the situation is understandable. This belief does not mean, however, that the outcome was divinely predetermined. It means only that the outcome was inevitable given the known antecedent conditions. In fact, the more predetermined the outcome was under the existing conditions and choices, the more likely changes in these conditions and choices would have led to a different outcome. In other words, the more a person is certain that outcome A was predictable under condition X, the more sure the person is that a change in condition X would have led to a change in outcome. Just as the person "knows" after the fact that bringing in the relief pitcher was bound to bring up the pinch hitter who would then hit the home run, the person "knows" equally well that leaving in the starting pitcher would have led to a strikeout and a different outcome.

Thus, the generation of counterfactual alternatives to reality might be enhanced by a hindsight bias through which one sees outcomes as inevitable under the preexisting set of conditions. These two kinds of biases might then go hand in hand and lead to a change in strategy in future situations. The fact is that the hindsight assessment and the inference about the counterfactual world are both likely to be incorrect. The outcome was not as predictable from antecedent conditions as people think (hindsight bias). Nor was an alternative outcome based on the mutation of some antecedent condition as likely as people think (counterfactual-

generation bias). Yet these two biases in judgment can combine, and they have the potential for turning good decisions into bad ones.

Foreseeability

The *foreseeability* of outcomes has played a role in several social-psychology theories. For example, any negative outcomes involved in a decision or a behavior are assumed not to contribute to cognitive dissonance when these negative outcomes were unforeseeable (Cooper & Fazio, 1984; Goethals, Cooper, & Naficy, 1979). Thus, finding out after the fact that a counterattitudinal essay will be used to change the attitudes of young people or will have significant negative personal ramifications should not contribute to feelings of cognitive dissonance or to attitude and behavior changes in the counterattitudinal direction.

Just as foreseeability is presumed necessary for the arousal of the aversive state of cognitive dissonance, it might likewise be proposed that the negative affective states and costly behaviors associated with counterfactual thinking will arise only if the negative outcome that is counterfactualized was foreseeable in light of antecedent conditions. Logically this relationship is true. Why should one feel blame or regret about an accident that occurred while driving home by a new route? Why should the father in Detroit suffer from so much guilt following a plane crash that was unforeseeable and of extremely low a priori likelihood? As we have shown, the affective experiences and the behaviors following counterfactual generation do not obey the dictates of logic and rationality. The accident that could have been avoided (even if unforeseeable) and the plane crash that could have been averted do indeed bring feelings of regret and guilt and self-directed anger. Apparently the arousal of these emotions and the instigation of different behaviors are sufficient to generate a counterfactual world that does not include the negative occurrence. Experiencing sadness or depression following an accident or the death of a loved one is, of course, logical. Experiencing guilt or regret or blame because the generation of a counterfactual world indicates that the negative event need not and (therefore) should not have occurred is quite another thing.

Interestingly, there may also be some question about whether foreseeability is necessary for the arousal of other aversive psychological states such as cognitive dissonance. Although the results of some studies do indicate the necessity of foreseeability (Cooper, 1971; Goethals et al., 1979), results of other research indicate that even unforeseen negative consequences can arouse dissonance. In Brehm's (1959) study, children wrote a counterattitudinal essay stating that they liked a vegetable that they

actually disliked. Subsequently, some participants learned of a totally unforeseen negative consequence: Their mothers would read these essays, and perhaps a childhood full of disliked vegetables would be in store. Those children in the unforeseen-negative-consequences condition showed added cognitive dissonance and attitude change. Sherman (1970) also examined the role of unforeseen negative consequences in attitude change. Participants wrote attitude-discrepant essays. Subsequently, some participants learned that their essays would appear as signed letters in their college newspaper, an unforeseen negative consequence. Compared to the essay writers who did not experience such consequences, these participants showed increased attitude change in the counterattitudinal direction. However, such enhanced change occurred only for those who had personal choice in whether or not to write the essay, indicating that initial dissonance arousal (which is based on personal choice) is necessary for unforeseen events to have their effects.

The point is that foreseeability is a concept defined by logic and objective standards. Even when conditions and events are in fact objectively unforeseeable, people seem to believe that they should have foreseen them. As long as a different decision could have undone the negative outcome, as long as a reasonable counterfactual can be generated, people may believe that they should have known this ahead of time and that they should have done things differently so that the counterfactual world would have taken place. Once again, the dysfunctionality of counterfactual thinking is apparent. Because the generation of alternative worlds is considered a "should have" in addition to a "could have," the negative affect of failing to act accordingly is then experienced, and costly changes in behavior may be initiated.

Self-Handicapping

Yet another area in which counterfactual generation can have behavioral and judgmental consequences is that of *self-handicapping*. Self-handicapping is essentially a strategy employed in situations that threaten self-esteem and in which future performance is uncertain (Berglas & Jones, 1978; for a review, see Arkin & Baumgardner, 1985). When faced with this uncertainty, many people choose to engage in behaviors that impede their ability to be successful at these tasks but that will serve as an excuse for failure. Thus, self-handicappers actually prepare for a counterfactual *in advance*, one that can relieve them of personal responsibility and feelings of poor performance by engaging in counterproductive behaviors before the event that will evoke the counterfactual after the fact (e.g., "If I had

studied the night before the exam instead of going out and drinking, then I would have passed").

Although there are no known experiments in which researchers have explored counterfactuals in the area of self-handicapping, the implications are clear. Engaging in behaviors ahead of time that will allow the generation of self-protective counterfactuals after the fact permits self-handicappers to maintain self-esteem in the face of uncertain conditions, even if their performance is bad. Although at first glance this use of counterfactuals seems beneficial, it is important to note that the self-handicapping activities themselves (e.g., taking drugs, not practicing, staying up too late) will probably lead to diminished performance regardless of the self-esteem benefits derived. Even if students do not feel bad or have lowered self-esteem after failing because they drank instead of studying, the fact is that their performance in all likelihood would have been better had they studied. Despite the maintenance of self-esteem that these counterfactuals provide, ultimately, engaging in the self-handicapping behavior hurts performance.

The idea that people can preconstruct counterfactual worlds to avert potential negative, undesirable states in the future is an interesting one. Consider the following situation. You are at the dog track. Before one of the races, you carefully consider all of the information available, and you decide that dog Number 3 in this race, with 5 : 1 odds, is the best bet of the day. You stand in line and are prepared to bet \$100 on dog Number 3 to win. The line moves very slowly, and just before you get to the betting window the bell sounds for the race. You are shut out. You cannot make your bet. You now watch the race. Which dog do you root for? Our own experience and discussion with others make us certain that you will root for any dog but Number 3. Why? The answer has to do with your preconstructions of the counterfactual that will be generated after the outcome is known. If dog Number 3 wins, the counterfactual will be something like "If only I had gotten in line 10 seconds earlier, I would now be \$500 richer." This counterfactual will bring much sadness and regret. If dog Number 3 loses, the likely counterfactual is "If I had gotten to bet, I'd now be \$100 poorer—this is my lucky day." Thus, the recognition ahead of time of the counterfactuals that are to come lead you to hope for the outcome that will be associated with the mood-enhancing, downward counterfactual. Moreover, this thinking process happens despite the fact that a win by dog Number 3 would be a strong indication of your decision-making skills. Once again it is clear how judgments and feelings depend as much on the alternatives to reality as on the reality itself. This idea that anticipatory regret can enter into judgments and preferences has been discussed by decision theorists (e.g., Bell, 1982, 1985; Loomes & Sugden, 1982; see also Gleicher et al., chapter 10; Miller & B. Taylor, chapter 11).

THE COSTLY MAINTENANCE OF DYSFUNCTIONAL BEHAVIORS

According to the previous analysis, counterfactual thinking can lead to the adoption of new behaviors or to changes in judgment strategies when, in fact, such changes are not rationally warranted and are not likely to lead to better outcomes in the future. There are other cases in which self-defeating behaviors may be maintained because of the ability to generate counterfactual worlds that support these behaviors.

Consider the compulsive gambler. The objective fact is that this person is consistently losing money. This fact should indicate poor decision making and an unfavorable situation and should lead to the conclusion that the person ought to stop gambling. Gamblers, however, have an interesting way of justifying their behaviors. They look at any particular gambling loss and see in it simple ways in which they could have easily won. If only a minor change had occurred, they would have won. Thus, they maintain the illusion that they are good decision makers and gamblers, that they will learn from this outcome, and that the future will be full of winning money. It is the generation of these close counterfactual wins that keeps gamblers gambling. Gilovich (1983) demonstrated this pattern in a striking study. In a college basketball game, participants had bet either on the University of California at Los Angeles (UCLA) or on the University of Louisville. During the game, an unexpected play occurred. A UCLA player missed an easy layup at a time when UCLA could have surpassed Louisville. Louisville ultimately won. In accounting for the outcome, only the losers (UCLA bettors) generated the counterfactual world in which the layup was made (in which case they would have won the bet). Thus, both winners and losers consequently believed that they were good decision makers, and their gambling continued.

In fact, those who operate gambling establishments understand very well the importance of close counterfactuals, cases in which a losing reality was within easy grasp of a winning outcome. Losers watching a photo-finish race will see in replay and in an enlarged picture just how close they were to winning. In the game of keno, not only do the lights under the winning numbers glow, but the lights also extend a little to all surrounding numbers, a practice that makes the close counterfactuals very accessible. Therefore, through their own motivational tendencies and with a little help from the gambling establishments, gamblers who lose can see how close to winning they really were and thus maintain their habitual gambling behaviors.

Consider a hustler in pool or tennis or any other game in which money might be at stake. The goal, of course, is not only to beat the opponent but to win in a particular way. It is important that the "mark," the one

duped, be left with a close counterfactual world in which winning was within grasp. A lucky drop of a ball off another ball in the final pool shot or a ball that barely clears the net in a close tennis match is sufficient to keep the mark playing again and again. Those close counterfactual worlds, which may be set up by the hustler or mentally manufactured by the mark, often keep people persisting in counterproductive behaviors.

Although one of the authors of this chapter has never been a tennis or pool hustler (although he may have been a mark), he *has* played his share of bridge. His favorite ploy is called the Grosvenor Gambit (Turner, 1980). You are declarer in a bridge contract, and you and your partner have bid to a slam that is unbeatable with simple play. Instead of simply playing the hand normally and claiming the win, you suddenly make a play that would allow the defenders to defeat the contract. However, in order to defeat the contract, they must assume that you have just made an unthinkable foolish play. They assume that you did not make such a gross error, they defend normally under the circumstances, and they hand the contract back to you. You then claim the contract, and they can then see how they could have easily defeated the contract. A normal result is achieved by an abnormal route, a situation affording the easy generation of a counterfactual world where a better outcome was achieved. This generation of a close counterfactual usually leads your opponents to scream at each other and at the same time to believe you are an idiot. You simply smile with a blank expression. They are yours for the rest of the evening.

CONCLUSION

Often automatically and sometimes inexplicably, people's minds wander to counterfactual alternatives to reality. Recent research and theorizing have been devoted to an understanding of when these counterfactual generations occur and what form and direction they will take. A key issue in this research and theorizing has been the identification of the beneficial functions served by these psychological mutations of reality (Markman et al., 1993; Roese, 1994). Moreover, several very important functions have been identified—affect regulation, preparation for the future, coping with negative life events, and bringing about perceptions of predictability and controllability for these events. Certainly these are important psychological functions.

More generally, it has been clear to social psychologists for a long time that reactions to life events, especially negative life events, depend on more than the objective reality of those events. It is the perception of the events and the interpretation of the meaning of those events that guide

humans' affective, cognitive, and behavioral reactions far more than the physical reality, or true meaning, of the events. In addition, social objects and events are rarely judged on an absolute basis. People require comparison objects against which to judge events and circumstances. These comparison objects usually take the form of other people who are in similar circumstances or of similar events that have occurred in the past. The meaning and evaluation of a person's experience depend very much on the outcome of such social comparisons. In cases in which there are no readily available comparison persons or available similar past life events, people may mentally manufacture such events. Furthermore, this function may be the most general and most important function of counterfactual generation—to provide the kind of social comparison that allows people to evaluate and interpret their life experience and to give meaning to that experience. Thus, the counterfactuals that people generate play an important role in guiding their perception of events and providing a comparison framework within which they can understand those events.

We certainly do not disagree that these functions are often served well by counterfactual thinking. Nor do we disagree with the conclusion that the potential benefits of counterfactual generation are many, in terms of improved affective state, a better understanding of circumstances and their causes, and a potential for improved performance and outcomes in the future. We ask only that the potential costs of counterfactual thinking also be considered—the ways in which mutations of reality can serve to induce negative affect and emotions, lead to biased and inappropriate judgments, and instigate counterproductive and self-defeating behaviors. We have tried to point to some caveats with respect to counterfactual thinking and to speculate about ways in which the generation of counterfactuals can hamper goal-directed behavior.

As with any simplifying principle, the generation and use of counterfactuals in judging circumstances and events can introduce characteristic errors and biases into the process. More than anything, counterfactual generation is a process that results in individuals' feeling that they have a better understanding of the causal structure of life's events. Yet, as we have shown, errors and biases based on counterfactual generation can creep into this causal understanding, resulting in unwarranted and debilitating negative emotions, irrational judgments and decisions, and behavioral changes that can be counterproductive. This tendency does not imply that people should try to eliminate counterfactual thinking from their lives or to desist from the mental construction of counterfactual alternatives to reality. It means only that psychologists should be more aware of the potential dysfunctional aspects of counterfactual generation and that they should recognize some of the errors and biases that can be associated with the process. At times, no doubt, people would be better

off not generating counterfactual alternatives to reality but rather simply dealing with that reality directly and accepting what comes along. Or in the words of Creedence Clearwater Revival's (1969) "Proud Mary": "And I never lost a minute of sleep, worryin' about the way things might have been."

REFERENCES

- American Law Institute (1985). *Model penal code: Official draft and explanatory notes*. Philadelphia: Author.
- Arkin, R. M., & Baumgardner, A. H. (1985). Self-handicapping. In J. H. Harvey & G. W. Weary (Eds.), *Attribution: Basic issues and applications* (pp. 169-202). San Diego, CA: Academic Press.
- Bargh, J. A., & Thein, R. D. (1985). Individual construct accessibility, person memory, and the recall-judgment link: The case of information overload. *Journal of Personality and Social Psychology*, 22, 293-311.
- Bargh, J. A., & Tota, M. E. (1988). Context-dependent automatic processing in depression: Accessibility of negative constructs with regard to self but not others. *Journal of Personality and Social Psychology*, 54, 925-939.
- Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, 54, 569-579.
- Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations Research*, 30, 961-981.
- Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations Research*, 33, 1-27.
- Berglas, S., & Jones, E. E. (1978). Drug choice as a self-handicapping strategy in response to noncontingent success. *Journal of Personality and Social Psychology*, 36, 405-417.
- Boninger, D. S., Gleicher, F., & Strathman, A. (1994). Counterfactual thinking: From what might have been to what may be. *Journal of Personality and Social Psychology*, 67, 297-307.
- Brehm, J. W. (1959). Increasing cognitive dissonance by a fait accompli. *Journal of Abnormal and Social Psychology*, 58, 379-382.
- Bulman, J. R., & Wortman, C. B. (1977). Attributions of blame and coping in the "real world": Severe accident victims react to their lot. *Journal of Personality and Social Psychology*, 35, 351-363.
- Clearwater, Creedence. (1969). Proud Mary. On *Bayou Country* [record]. Berkeley, CA: Jondora Music-BMI.
- Cooper, J. (1971). Personal responsibility and dissonance: The role of foreseen consequences. *Journal of Personality and Social Psychology*, 18, 354-363.
- Cooper, J., & Fazio, R. H. (1984). A new look at dissonance theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17, pp. 229-266). New York: Academic Press.
- Davis, C. G., Lehman, D. R., Silver, R. C., Wortman, C. B., & Ellard, J. H. (1994). *Taking more blame than is one's due: The role of counterfactual thinking*. Unpublished manuscript.
- Davis, C. G., Lehman, D. R., Wortman, C. B., Silver, R. C., & Thompson, S. C. (1995). The undoing of traumatic life events. *Personality and Social Psychology Bulletin*, 21, 109-124.
- Fischhoff, B. (1975). Hindsight-foresight: The effects of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 288-299.
- Gavanski, I., & Wells, G. L. (1989). Counterfactual processing of normal and exceptional events. *Journal of Experimental Social Psychology*, 25, 314-325.
7. DYSFUNCTIONAL IMPLICATIONS
- Gilovich, T. (1983). Biased evaluation and persistence in gambling. *Journal of Personality and Social Psychology*, 44, 1110-1126.
- Gleicher, F., Kost, K. A., Baker, S. M., Strathman, A. J., Richman, S. A., & Sherman, S. J. (1990). The role of counterfactual thinking in judgments of affect. *Personality and Social Psychology Bulletin*, 16, 285-295.
- Goethals, G. R., Cooper, J., & Naficy, A. (1979). Role of foreseen, foreseeable, and unforeseeable behavioral consequences in the arousal of cognitive dissonance. *Journal of Personality and Social Psychology*, 37, 1179-1185.
- Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgment of past events after the outcomes are known. *Psychological Bulletin*, 107, 311-327.
- Higgins, E. T., King, G. A., & Mavin, G. H. (1982). Individual construct accessibility and subjective impressions and recall. *Journal of Personality and Social Psychology*, 43, 35-47.
- Hilton, D. J., & Slugoski, B. R. (1986). Knowledge-based causal attribution: The abnormal conditions focus model. *Psychological Review*, 93, 75-88.
- Janoff-Bulman, R. (1979). Characterological versus behavioral self-blame: Inquiries into depression and rape. *Journal of Personality and Social Psychology*, 37, 1798-1809.
- Johnson, J. T. (1986). The knowledge of what might have been: Affective and attributional consequences of near outcomes. *Personality and Social Psychology Bulletin*, 12, 51-62.
- Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 93, 136-153.
- Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201-208). New York: Cambridge University Press.
- Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, 94, 211-228.
- Kruglanski, A. W., & Freund, T. (1983). The freezing and unfreezing of lay inferences: Effects on impression primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology*, 19, 448-468.
- Landman, J. (1987). Regret and elation following action and inaction: Affective responses to positive versus negative outcomes. *Personality and Social Psychology Bulletin*, 13, 524-536.
- Landman, J. (1993). *Regret: The persistence of the possible*. New York: Oxford University Press.
- Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, 32, 311-328.
- Levine, M. (1959). A model of hypothesis behavior in discrimination learning set. *Psychological Review*, 66, 353-366.
- Lipe, M. G. (1991). Counterfactual reasoning as a framework for attribution theories. *Psychological Bulletin*, 109, 456-471.
- Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal*, 92, 805-824.
- Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology*, 47, 1231-1243.
- Markman, K. D., Gavanski, I., Sherman, S. J., & McMullen, M. N. (1993). The mental simulation of better and worse possible worlds. *Journal of Experimental Social Psychology*, 29, 87-109.
- Markman, K. D., Gavanski, I., Sherman, S. J., & McMullen, M. N. (1995). The impact of perceived control on the imagination of better and worse possible worlds. *Personality and Social Psychology Bulletin*, 21, 588-595.
- Martin, L. L., & Tesser, A. (1989). Toward a motivational and structural theory of ruminative thought. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 306-326). New York: Guilford.

- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207-238.
- Medvec, V. H., & Gilovich, T. (in press). When less is more: Counterfactual thinking and satisfaction among Olympic medalists. *Journal of Personality and Social Psychology*.
- Miller, D. T., & Gunasegaram, S. (1990). Temporal order and the perceived mutability of events: Implications for blame assignment. *Journal of Personality and Social Psychology*, *59*, 1111-1118.
- Miller, D. T., & McFarland, C. (1986). Counterfactual thinking and victim compensation: A test of norm theory. *Personality and Social Psychology Bulletin*, *12*, 513-519.
- Miller, D. T., Taylor, B., & Buck, M. L. (1991). Gender gaps: Who needs to be explained? *Journal of Personality and Social Psychology*, *61*, 5-12.
- Miller, D. T., & Turnbull, W. (1990). The counterfactual fallacy: Confusing what might have been with what ought to have been. *Social Justice Research*, *4*, 1-19.
- Miller, D. T., Turnbull, W., & McFarland, C. (1989). When a coincidence is suspicious: The role of mental simulation. *Journal of Personality and Social Psychology*, *57*, 581-589.
- Miller, D. T., Turnbull, W., & McFarland, C. (1990). Counterfactual thinking and social perception: Thinking about what might have been. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 305-331). New York: Academic Press.
- Niedenthal, P. M., Tangney, J. P., & Gavanski, I. (1994). "If only I weren't" versus "if only I hadn't": Distinguishing shame and guilt in counterfactual thinking. *Journal of Personality and Social Psychology*, *67*, 585-595.
- Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 87-108.
- Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology*, *46*, 69-81.
- Roese, N. J. (1994). The functional basis of counterfactual thinking. *Journal of Personality and Social Psychology*, *66*, 805-818.
- Roese, N. J., & Olson, J. M. (1993a). Self-esteem and counterfactual thinking. *Journal of Personality and Social Psychology*, *65*, 199-206.
- Roese, N. J., & Olson, J. M. (1993b). The structure of counterfactual thought. *Personality and Social Psychology Bulletin*, *19*, 312-319.
- Roese, N. J., & Olson, J. M. (1994). *Counterfactuals, causal attributions, and the hindsight bias: A conceptual integration*. Unpublished manuscript.
- Roese, N. J., & Olson, J. M. (1995). Outcome controllability and counterfactual thinking. *Personality and Social Psychology Bulletin*, *21*, 620-628.
- Sears, D. O. (1983). The person-positivity bias. *Journal of Personality and Social Psychology*, *44*, 233-240.
- Sherman, S. J. (1970). Attitudinal effects of unforeseen consequences. *Journal of Personality and Social Psychology*, *16*, 510-520.
- Sherman, S. J. (1980). On the self-crasping nature of errors of prediction. *Journal of Personality and Social Psychology*, *39*, 211-221.
- Skov, R. B., & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology*, *22*, 93-121.
- Smith, E. R., & Zarate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review*, *99*, 3-21.
- Taylor, S. E., Lichtman, R. R., & Wood, J. V. (1984). Attributions, beliefs about control, and adjustment to breast cancer. *Journal of Personality and Social Psychology*, *46*, 489-502.
7. DYSFUNCTIONAL IMPLICATIONS
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381-403). New York: Academic Press.
- Turnbull, W., & Mawhinney, M. (1986). *Counterfactual thinking and self-blame*. Unpublished manuscript.
- Turner, F. B. (1980). The Grosvenor Gambit. In E. Kaplan & J. Rubens (Eds.), *Bridge World humor* (pp. 38-42). New York: Bridge World Magazine.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207-232.
- Walster, E. (1966). Assignment of responsibility for an accident. *Journal of Personality and Social Psychology*, *3*, 73-79.
- Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, *23*, 273-281.
- Wells, G. L., & Gavanski, I. (1989). Mental simulation of causality. *Journal of Personality and Social Psychology*, *56*, 161-169.
- Wells, G. L., Taylor, B. R., & Turtle, J. W. (1987). The undoing of scenarios. *Journal of Personality and Social Psychology*, *53*, 421-430.
- Wood, J. V., Saltzberg, J. A., Neale, J. M., Stone, A. A., & Rachmiel, T. B. (1990). Self-focused attention, coping responses, and distressed mood in everyday life. *Journal of Personality and Social Psychology*, *58*, 1027-1036.