APPLIED COGNITIVE PSYCHOLOGY, VOL. 10, S113–S124 (1996)

The Role of Counterfactual Thinking in Reasoning

STEVEN J. SHERMAN
Indiana University

ALLEN R. McCONNELL
The Pennsylvania State University

SUMMARY

Counterfactual generation is an important part of reasoning. Both the judgment of events and affective reactions to those events depend not only on the events themselves, but on counterfactual alternatives to those events. Counterfactual thinking serves several positive functions. However, there are also dysfunctional aspects. First, judgments of general versus specific instances are often inconsistent, and this leads to problematic, irrational decisions. We explain these inconsistencies by suggesting that specific instances easily afford counterfactuals, and are judged in the context of these counterfactuals. Alternatively, general cases are evaluated in terms of quite different contrast cases, global expectations. Second, in assigning blame for the negative outcome of a chain of events, people assign too much causality to recent events. Our explanation is that these recent events are most accessible and are most likely to be mutated in the course of counterfactual generation. Such mutability is important in causal assignment.

People have a great deal of trouble accepting the reality of their lives and being satisfied with this reality. Instead, they have a compelling propensity to mentally alter life's circumstances and to reflect on 'what could have been' or 'what should have been'. People create such an alternative reality by mutating certain facts, events and circumstances that were antecedent to their current situation. This obsession with creating and imagining other possible worlds is known as counterfactual thinking. Counterfactual thinking has become a growing area of speculation and research in psychology since a seminal paper by Kahneman and Miller in 1986.

The generation of counterfactual worlds is important because the imagination and consideration of what could have been and the realization that current circumstances were not inevitable have significant affective, behavioural and judgmental consequences. Our reasoning and judgments about current situations and outcomes are determined not only by the value and the facts of those situations but also by the outcomes that did not occur, but might have occurred, if only certain antecedent events had been slightly different.

A collection of chapters by Roese and Olson (1995) provides an excellent summary of recent work on counterfactual generation. This research has focused primarily on

Address correspondence to: Steven J. Sherman, Psychology Department, Indiana University, Bloomington, IN 47405, USA.

CCC 0888–4080/96/SI0113–12

Accepted 2 July 1996

two issues. First, which aspects of reality are most mutable? That is, of all the possible alternative worlds, which counterfactuals are most likely to be mentally constructed? Second, what are the affective consequences of counterfactual generation in terms of regret, satisfaction and guilt?

Work in our own laboratory has focused on a somewhat different question: What are the functions of counterfactual thinking? Given their prevalence and their seemingly automatic and inescapable generation, counterfactuals must indeed serve important psychological functions. In our work (Markman, Gavanski, Sherman and McMullen, 1993, 1995) and in the work of others (Roese, 1994), three such functions of counterfactual thinking have been specified. These functions are all beneficial and adaptive.

(1) *Affect regulation.* The imagination of other possible worlds can help people feel better about their current reality. This function is generally served by downward counterfactuals, i.e., alternatives that change reality to a worse one. For example, consider an automobile accident where there is $5,000 worth of damage to the car, and the driver suffers a broken leg. An example of a downward counterfactual might be: 'If I hadn't fastened my seat belt 5 minutes before the accident, I would have been killed'. Despite the objectively negative aspects of life's circumstances, the realization that these circumstances could have been even worse if certain small changes in antecedent conditions had occurred makes current reality more acceptable. The easier it is to generate downward counterfactuals, the greater the regulation of negative affect. The idea that objects and events are judged relative to alternative possibilities is, of course, consistent with many approaches to reasoning and judgment (Schwarz and Bless, 1992; Thibaut and Kelley, 1959).

(2) *Preparation for the future.* Counterfactuals are most often generated in the case of negative life events and circumstances (Gleicher, Kost, Baker, Strathman, Richman and Sherman, 1990; Landman, 1987). When current reality is difficult and when bad things happen, people dwell on how life might have been better 'if only . . .'. Such upward counterfactual generations (alternatives that change reality into a better world) clearly have the potential for worsening one's affective experience. A current bad reality is judged as *even worse* in light of the more positive alternatives that might have come about. On the other hand, if life could have been even better given the imagined mutations, by extension life can be and will be better in the future if these mutations are now actually made. Thus, upward counterfactual generation gives one hope for the future and prepares one for the future (Markman *et al.*, 1993; Roese, 1994). For example, again considering the automobile accident, an upward counterfactual might be: 'If only I had come to a complete stop at the stop sign, there wouldn't have been an accident'. Such upward counterfactuals might leave one feeling dissatisfied and blameworthy because current reality *could have been* avoided, but at least one is better prepared for future circumstances and outcomes are likely to be better next time.

(3) *Feelings of controllability.* Control over one's life is a key motivation. Many counterfactuals indicate to the person involved how *they* could have prevented a current bad reality. At least they had control (and will in the future), and whatever happened was not simply thrust upon them or a matter of whimsical fate. Even rape victims and cancer patients generate counterfactuals in which *they* could have prevented the rape or disease by altering some aspect of their own behaviour (Janoff-Bulman, 1979; Taylor, Lichtman and Wood, 1984). On the one hand, such counterfactuals can have negative affective consequences because one now feels responsible for negative life events. On the other hand, the feeling of control over the past, and thus the feeling of control in the future, may more than compensate for those increases in negative affect.

In summary, counterfactual generation serves important positive psychological functions, so it is no wonder that counterfactuals are so ubiquitous. More recently, however, we have been focusing on some of the unintended dysfunctional aspects of counterfactual generation as well (Sherman and McConnell, 1995). These dysfunctional aspects influence emotional states, behaviour and reasoning. In the present paper, we focus on two of the most important and interesting aspects of counterfactual thought that can be dysfunctional because they lead to biased, inconsistent, or suboptimal reasoning and judgment.

INCONSISTENT JUDGMENTS OF GENERAL VERSUS SPECIFIC INSTANCES

Recently we have been intrigued by a number of anecdotal cases in which judgments appeared to be irrational, inconsistent, or at least difficult to understand. These cases all involve judgments and decisions that people make about general or abstract instances, on the one hand, versus specific events or occurrences on the other hand. Some examples of such judgments appear to be quite puzzling as well as potentially dysfunctional for rational decision-making.

In the summer before the 1992 Presidential election, and prior to the Democratic Convention, it was clear that George Bush was the Republican nominee, but it was unclear who the Democratic nominee would be. Eight Democratic contenders were in the running. Many polls were taken. In these polls, Bush lost consistently to an abstract opponent identified only as 'the Democratic nominee'. On the other hand, in the polls Bush beat each and every viable specific Democratic candidate. There are other similar examples that illustrate the power of specific instances. A specific little girl who fell down a well in Texas was sent a million dollars in contributions. However, relatively little is given through charities to 'needy children' in general. Hundreds of thousands of dollars were authorized, with widespread public support, to save two specific whales that were trapped in icy seas, but relatively little money is offered by governmental agencies or by individuals for a general 'save the whale' campaign.

The person positivity bias (Sears, 1983) represents another case where evaluations of general versus specific cases are not consistent. Sears found that individuals are each judged as far more positive than is the general group to which all of them belong. Thus, Congress in general is rated very negatively, and yet each and every member of Congress is evaluated quite positively when judged in isolation.

In accounting for these cases of disjuncture between judgments of general versus specific cases, Sherman, Beike and Ryalls (1995) have proposed several mechanisms. These include mechanisms of integration of information from parts to a whole and

of degradation of information from a whole to its parts. The accounts also include differences in cognitive representation of general versus specific information, as well as several dual-processing accounts (e.g., on-line vs. memory-based processing, central vs. peripheral processing, and rational vs. experiential processing).

One other account is most germane to the current paper, and involves counterfactual generation. This account is most applicable to cases where there is an inconsistency of reactions to a general failure rate as opposed to reactions to any specific instance of failure, even when some number of failures is to be expected. For example, it is understood and it is acceptable that a certain percentage of patients will die during surgeries. It is even understood that medical judgment is fallible and that a certain number of suboptimal judgments about medical cases will be made by doctors. Yet this willingness to accept errors at the general level is not matched by an acceptance of the death of any specific patient or any specific suboptimal judgment about an individual medical case. Blame and law suits in specific instances replace the calm acceptance of a particular judgmental error rate at the general level.

Nowhere are such discrepancies in the judgment of general error rates versus specific errors more apparent than in the sports world. A college basketball coach and team fans are happy to accept a general turnover rate of ten per game. And yet any specific turnover is met with moaning, booing, screaming, chair throwing and worse. Specific errors are not simply accepted as a part of a reasonable error rate, even when the number of specific errors is well within such a general error rate. Such inconsistencies in judgment can be quite problematic for rational decision-making.

In understanding this difference in reactions to general error rates versus specific errors, Sherman *et al.* (1995) suggested that, when people are faced with failures or errors, they engage in a causal analysis to understand and evaluate the negative outcomes. As they try to understand, judge, or provide causal explanations for these failures, they must often depend on a comparison case as a standard or context of judgment. Which comparison case is adopted will very much affect the judgment or interpretation of the event itself (McGill, 1989; McGill and Klein, 1993). Consider the case of a man who has an ulcer attack. In asking 'Why?' the man's wife might answer 'Because of the spicy Hunan meal that he ate'. Her comparison case involves other days when the husband did not have an ulcer attack. What was different on this day was the spicy meal. On the other hand, the doctor's answer to the same question might be 'Because of a weak stomach wall'. Her comparison case involves other patients who do not have ulcer attacks. What is different about this man from the others is his weak stomach wall.

One important kind of comparison case for failures involves prior expectancies, the anticipation of what is likely to happen. A team that wins 65% of its games is judged as having a very successful season if the expectation was for a losing year, but is judged as having an awful season if the prior expectancy was for an undefeated season. We would argue that prior expectancies are likely to be used as a standard of comparison for general instances involving failures or error rates. If the global error rate is within one's prior expectancy, then all is well. Only if the error rate exceeds the prior expectancy is there dissatisfaction or anger.

Specific events, however, are not embedded in this same kind of understanding. Instead, each specific event (e.g., an error or a turnover) may bring its own comparison standard into being after the fact. This is how Kahneman and Miller (1986) conceived of counterfactual generations—post-event constructed alternative

cases that are generated after an event occurs. This mutated reality is then used as a comparison case for the factual world. Specific errors, with their identifiable and unique features, easily afford such post-event computed comparison cases. The mutuation of unusual, unexpected, or salient features of the specific error or turnover yields an upward counterfactual in which the specific error need not have occurred 'if only . . .'. What *could* have been is often falsely taken as what *should* have been (Miller and Turnbull, 1990), which produces dysfunctional affective reactions and judgments. The reaction is likely to be blame, anger and dissatisfaction—the typical affective responses when upward counterfactuals are generated (Boninger, Gleicher and Strathman, 1994; Markman *et al.*, 1993). These affective reactions are likely to be translated into unreasonable decisions. For example, coaches may change a generally good strategy into a bad one because of one specific failure (Baron and Hershey, 1988). In addition, coaches get blamed by fans and owners for single, specific failures, and they may lose their jobs. Players get blamed by their coaches for specific errors, and might get benched or traded. Very few fans, managers and coaches seem immune to these dysfunctional aspects of counterfactual generation for specific errors or failures. Tony LaRussa, the former manager of the Oakland As (and now of the St. Louis Cardinals), for example, seems immune to emotional reactions to specific errors and failures. Perhaps his great success is due to this ability.

General error rates or failure rates, on the other hand, have no particular features that can be mutated into a counterfactual alternative. Thus, one is more likely to use pre-existing pre-event expectancies as contrast cases. If the general error rate is consistent with the expectancy, then it is acceptable. Only if the general error rate is higher than the expectancy will it lead to unhappiness. This, then, can explain why an overall error rate may be acceptable, and yet any specific error will not be acceptable. It is the use of different contrast cases (i.e., counterfactual alternatives for specific instances and pre-existing expectancies for general instances) that leads to differences in reactions to and differences in judgments about general versus specific failures.

When looking backward in time, specific events easily afford counterfactuals by mutating past events that serve as a standard of comparison for the present. General cases do not afford counterfactuals and are instead compared to prior expectancy levels. Interestingly, just as specific events easily yield specific alternative pasts, so do they easily yield visualizable and detailed future scenarios. For example, it is easy to visualize the little girl trapped in the well who plays happily at home once she is saved. It is easy to image the two trapped whales swimming happily with their pod once they are free. It is far more difficult to visualize the outcomes of general cases (e.g., giving money in order to 'reduce world hunger'). One cannot clearly image less hunger in the world or fewer deaths from AIDS. But any specific case can easily be imaged forward in time in a clear scenario. Thus, the visualization of alternatives, whether counterfactuals of past events or possible worlds of the future, impact on judgments and affect them.

Some years ago, Sherman, Zehner, Johnson and Hirt (1983) considered the impact of imaging and explaining hypothetical future events on judgments of the probabilities of those events. For example, subjects were asked to think about an actual upcoming football game, were given real facts about the teams, and were asked to imagine and explain why one (or the other) team managed to win easily.

After generating such hypothetical future events, subjects were asked for their actual judgments of how the game would turn out. The generation of hypothetical outcomes greatly affected subjects' judgments in the direction of the explained hypothetical outcome. Interestingly, some subjects generated very specific scenarios (e.g., 'The quarterback throws a last second pass to the wide receiver, and the crowd cheers as he crosses the goal line.'). Other subjects explained the outcome by citing general factors (e.g., 'UCLA is a better overall team.'). The generation of specific scenarios as opposed to using more general factors had the greatest impact on subsequent judgments of the probability of victory and they were the most resistant to change from new information indicating that their favored team might lose.

With regard to current concerns, we would suggest that it is easy to generate a clear, visualizable scenario for a specific upcoming event. And such a generated scenario ought to have an impact on judgments of what actually will occur such that the event imagined appears more likely. In the case of general events, no such visualizable scenario is possible, and thus no large effect on judgments about the future is expected. Thus, asking subjects to generate a scenario for a specific upcoming game should very much affect judgments of that game. However, asking subjects to imagine and explain why a team will have a great overall season should lead to the use of global facts and statistics in order to explain such an event and should not greatly impact judgments of how the team's season will actually go.

In summary, judgments of general versus specific cases are often inconsistent and seemingly irrational. Many mechanisms have been suggested to account for such inconsistencies (Sherman *et al.*, 1995). One of the prominent explanations is in terms of counterfactual generation. Only specific cases afford counterfactuals, and thus only specific cases are subject to the affective and judgmental consequences of counterfactual thinking. General cases, on the other hand, are evaluated in terms of global expectancies rather than counterfactual worlds.

JUDGMENTS OF CAUSALITY BASED ON THE TEMPORAL ORDER TO EVENTS

People must often judge the causality of an outcome that occurs after a long and complex chain of events. In this section, we consider only the case of a series of events that are not causally related to each other. In such chains, each event does not cause the next event, but rather these independent events happen to take place in a particular temporal sequence (for a discussion of causally related temporal events, see Johnson, Ogawa, Delforge and Early, 1989; and Miller, Turnbull and McFarland, 1990).

It is quite clear that counterfactual thinking is central to attributions of causality (McGill and Klein, 1993; Wells and Gavanski, 1989). That which is most easily mutated in order to create a counterfactual is perceived as the cause of the factual event. Thus, consider the man who spends his usual hour and a half for lunch and then chooses to drive on a new route to the airport. There is construction along this new route, and he is delayed by 20 minutes. He misses his flight by 5 minutes. The cause of his missing the plane is likely to be seen as his having taken a new route rather than the excessive time that he took with lunch. This is because the new route is the most non-normative (and thus, mutable) aspect of the scenario. As past

research has demonstrated (e.g., Wells and Gavanski, 1989), mutability and causality are intimately related. How does this fact help us to understand the allocation of causality between two (or more) causally relevant factors that occur in temporal sequence prior to an outcome?

Miller and Gunasegaram (1990) found that events that occur later in a temporal chain are given greater emphasis in assessments of causality for an outcome than earlier events in the same non-causal sequence. In one of their experiments, participants who imagined themselves performing poorly on a test 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, they blamed the teacher's poor exam instead of their own study habits for failure. On the other hand, when participants were told that the test was constructed prior to their studying, they blamed their poor study habits rather than the teacher's unfairness as the cause of their poor performance. Miller and Gunasegaram (1990) explained these results by proposing that later events in a temporal chain are easier to mutate than earlier events. Early events are taken as background or givens, and are therefore immutable. The mutability of later events is why they are judged as causal for the outcome. Because the two events (exam construction and studying) are independent of each other and because the temporal sequence is basically a chance occurrence, these judgments of causality are irrational and unwarranted. Yet they may have subsequent dysfunctional consequences if students base their future study habits or teacher evaluations on these counterfactual-driven judgments of causality.

This notion that people will mutate the last event in a temporal chain and assign causality to this event received further support in another demonstration provided by Miller and Gunasegaram (1990). They presented the following situation to participants. Jones and Cooper flip a coin separately. If their coins match (i.e., both heads, or both tails), they each win $1000. If the two coins do not match, neither individual wins anything. Jones goes first and tosses a head. Cooper goes next and tosses a tail. As a result, neither person wins anything. The question posed to participants was who would experience more blame: Cooper or Jones? Despite the fact that the two events were independent, 92% of participants predicted that Jones would blame Cooper more than vice versa. In other words, they put the blame on the person at the end of the temporal chain despite the fact that they could have just as easily concluded that Cooper and Jones would have won the money, 'if only Jones had tossed a tail'. Thus, it does not seem rational to blame Cooper when Jones's action contributed equally to their not winning the money.

Based on the findings of Miller and Gunasegaram, it might seem appropriate to conclude that people will mutate the last event in a temporal chain and judge it as causal. In the case of independent events (where prior events are not causes of later events), the most recently encountered events (i.e., those at the end of the temporal chain) will be most available in memory and thus likely candidates for mutation (Kahneman and Miller, 1986).

However, this counterfactually driven explanation has recently been challenged by Spellman (1995), who argues that such judgments are computed by using probabilistic assessment rather than counterfactual generation. According to Spellman's 'crediting causality' hypothesis, perceivers rely on several probabilistic indicators when assessing which of two independent events is causal for an

outcome's occurrence: the initial probability of the outcome occurring, the change in probability of the outcome occurring after the first but before the second event occurs, and the change in the probability of the outcome occurring by adding the second event to the first event. The event that more significantly alters the likelihood of the outcome in the direction of the actual outcome is credited with the greater causality. In the Jones–Cooper example, the prior probability of winning is 0.5. After Jones's toss of heads, the probability of winning is still 0.5—no change. It is Cooper's toss that reduces the probability from 0.5 to 0. Thus, Cooper is given the lion's share of the blame. In a series of studies, Spellman used scenarios similar to the Jones and Cooper coin flip case, but more complicated, in order to explore how the order of independent events influences causal judgments. In general, she found that participants appeared to recompute the likelihood of the outcome (i.e., Jones and Cooper not winning the money) for the second event after assessing the probability of the outcome given the first event. As such, Spellman claims that probabilistic assessments, not counterfactual generation, can explain these causal judgments.

Therefore, the question remains: Do counterfactuals play a role in judgments of causality in temporal chains? Although Miller and Gunasegaram (1990) claimed that counterfactual thinking is responsible for such judgmental biases, they provided no evidence of the mediating role of counterfactuals. As is the case of many studies, counterfactuals are often considered only as dependent variables rather than as mediators (there are exceptions, e.g., Davis, Lehman, Silver, Wortman and Ellard, 1996). In a recent series of studies, we explored the potential mediational role of counterfactual thinking in dysfunctional judgments that may occur when people systematically mutate the last event in a temporal chain.

In one experiment, we examined how the last game of the season may be overemphasized in judgments of the overall performance of sports teams. Although many sports fans say that 'every game counts' in their personal assessments of sports teams, they sometimes admit (albeit quietly at times) that they are overly influenced by games at the end of the season. Why is this the case? A counterfactual explanation would posit that the last game will be more mutable because of its salience. Accordingly, the last game will be viewed as especially causal of the entire season record, despite the relative independence of game outcomes.

To test this idea, we presented 93 undergraduates with a simulated 10-game season in basketball. Six teams played twice against all five other teams during the course of the season. At the end of each week, participants rated the quality of each team's performance *for the entire season* using a poll ranking system (the best team is ranked 1, the second best is ranked 2, etc.). At the end of the season, two teams (the Blue Team and the Red Team) had a record of nine wins and one loss, each beating the other four teams twice (eight of their nine wins) by equivalent amounts during the season. In week 5, the Blue Team beat the Red Team. In week 10, the Red Team beat the Blue Team by six points.

One question that we pursued was whether participants would still give greater weight to the last event (week 10) even when the earlier event (the blue team winning in week 5) was a more lopsided victory than the later event when the red team won. Thus, we manipulated how badly the Blue Team beat the Red Team in week 5: either by 6 points (same as the Red Team's margin of victory in week 10), 12 points (thus the Blue Team, objectively, played better than the Red Team), or 24 points (thus the Blue Team's overall performance was far superior to the Red Team's). Although we

anticipated that at some point (e.g., in the 24-point condition) participants would concede that the Blue Team should be number 1 (despite losing the last game), we were surprised that the manipulation of the week 5 outcome did not affect their final rankings. Regardless of experimental condition, the majority of subjects ranked the Red Team ahead of the Blue Team. Thus, participants relied heavily on the last game for their final rankings, even though they were told explicitly that their rankings should reflect the entire season.

We were also intrigued by some of our internal analyses. In addition to providing weekly rankings, participants were asked at the end of the experiment to generate counterfactuals about what could have occurred differently to produce different final rankings. Those who spontaneously evoked the 'if only the Blue Team had won the last game' counterfactual were especially likely to rate the Red Team ahead of the Blue Team in the final rankings ($r=0.25$, $p<0.02$). Moreover, those who generated the counterfactual 'if only the Red Team had won the week 5 game' were more likely to rank the Blue Team ahead of the Red Team (i.e., not fall prey to placing the majority of the blame on the last event; $r=0.33$, $p<0.01$). So although these data suggest that counterfactuals are generally related to the final judgments rendered by participants, it appears that it is difficult to overcome the overutilization of the last event in judgment. It should not be surprising, then, that national polls also give greater weight to late-season losses than to early-season losses.

In a second experiment, we also used a basketball theme, but this time used assessments of player performance instead of team rankings. We utilized the classic case of the basketball player who, at the end of the game, misses the potential winning basket. This player is often 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. One would expect that the ease of mutating the last event in the temporal chain will lead to placing the blame on the person who misses at the end when others might be as much, or even more, responsible for the ultimate outcome. The generation of such counterfactuals might also carry with it behavioural implications, such as influencing who will start the next game. Further, if counterfactuals do play a role in judgments, we might observe that they mediate such judgments.

Eighty-three undergraduates at a computer workstation read a text-based transcript that simulated the last 15 minutes of a basketball game. Two teams played each other. Each team had two players. At the end of the game, the Green Team was losing by one point and a player from the Green Team had the opportunity to shoot two foul shots with essentially no time remaining in the game. If this player (hereafter, Loser) made both foul shots, the Green Team would win the game. However, if Loser made only one of the two shots, the teams would end up tied. And, if Loser missed both shots, the Green Team would lose the game. Loser missed both shots. After the conclusion of the game, participants generated counterfactuals about how the game could have turned out differently, and they rated each player's performance (each on 9-point scales). They were also asked which of the two Green Team players should get more future playing time in an upcoming game.

The critical (between-subjects) manipulation in this experiment was the performance of Loser's team-mate (hereafter, Team-mate) earlier in the game. In one condition, Team-mate also missed two foul shots (0–2 condition; making the Loser's and the Team-mate's performances identical). In a second condition, Team-mate missed four foul shots (0–4 condition; objectively worse than the Loser's

performance). Finally, in a third condition, Team-mate missed eight foul shots (0–8 condition; far worse than the Loser's performance). Across the whole game segment, Loser and Team-mate had identical performances, except in terms of foul shooting in the 0–4 and 0–8 conditions (where the Team-mate performed more poorly than the Loser).

The first question of interest was whether or not participants produced the counterfactual, 'If only the Loser had made those last foul shots' in order to change the game's outcome. Analyses revealed that 43% of participants in the 0–2 condition generated a counterfactual about the Loser's last foul shots, whereas 26% did in the 0–4 condition and 22% did in the 0–8 condition. Although there was a tendency to form this counterfactual in all three conditions, participants produced this counterfactual more often in the 0–2 condition than in the 0–8 condition [$\chi^2(2)=6.62$, $p<0.04$]. The second measure of interest, the rating of the Loser's quality of play relative to the Team-mate's quality of play, revealed that the Team-mate was judged to have played better in the 0–2 and 0–4 conditions, but that the Loser was judged to have played better in the 0–8 condition [$F(2, 79)=12.72$, $p<0.001$]. Thus, participants rated the Team-mate's performance as superior to the Loser not only in conditions where the two players' performance was identical (i.e., the 0–2 condition), but even in cases where the Team-mate's performance was objectively inferior to the Loser (i.e., the 0–4 condition).

Also of interest was the judgment of who should play more in the future. Perceivers who do not weigh the last shots more heavily than other shots in their judgments should express no preference in the 0–2 condition, but prefer the Loser in the 0–4 and 0–8 conditions. However, the data revealed that participants preferred playing the Team-mate (relative to the Loser) in both the 0–2 condition (71% preferred the Team-mate and 29% preferred the Loser despite identical objective performance) and the 0–4 condition (67% preferred the Team-mate and 33% preferred the Loser even though the Loser played better). Only in the 0–8 condition (where the Team-mate's performance was far worse than the Loser's) did participants express a preference for the Loser (74% vs. 26% for the Team-mate). In all three conditions, these preferences differed significantly from chance (binomial test, $p<0.06$), and there was a strong effect of experimental condition, $\chi^2(2)=13.75$, $p<0.001$. In sum, participants were strongly influenced by the Loser's missing the last foul shots even when the Team-mate's performance was objectively worse. Only in extreme cases (the 0–8 condition) was this pattern different.

In addition to these between-condition findings, correlational analyses were conducted to explore the mediating role of counterfactuals in these judgments. It was found that the likelihood of generating the counterfactual 'if only the Loser had not missed the last foul shots' was positively related to the preference to play the Team-mate in the future game ($r=0.32$, $p<0.01$). This relationship, coupled with the fact that the experimental condition was predictive of both the likelihood of generating the 'blame the Loser' counterfactual and the future playing time decision, permitted us to conduct a mediational analysis to see if the counterfactual itself played a causal role in the decision to bench the Loser as a function of experimental condition (for details concerning mediational analyses, see Judd and Kenny, 1981). This analysis revealed that the generation of the 'blame the Loser' counterfactual partially mediated the relationship between experimental condition and future playing time for the Loser ($r=0.37$, $p<0.01$, without the mediator; $r=0.23$, $p<0.05$, with the mediator entered first). So although complete mediation was not observed, this partial mediation provides good evidence that counterfactual thinking does play a causal role in shaping our judgments of temporal events.

CONCLUSION

It is clear that reactions to life events depend on more than the objective reality of those events, but also depend on the interpretation of those events in light of standards of comparison and contrast cases. It is also clear that counterfactual alternatives to reality often serve as such contrast cases. The generation of worlds that could have been can help people get through difficult life experiences by mitigating negative affect, by giving them hope for the future, and by allowing them to perceive a sense of control over their circumstances.

What has been less clear is that there is also a down side, a costly side, to counterfactual thinking. We have demonstrated these dysfunctional outcomes in two types of judgments, the consideration of general versus specific cases and the analysis of causality for temporal sequences. In these instances, mutations of reality can induce negative affect and emotions, can lead to biased reasoning and inappropriate judgments, and can instigate counterproductive and self-defeating behaviours. This potential for dysfunctional consequences does not imply that people should try to eliminate counterfactual thinking from their mental lives. It simply means that we should be more aware of the errors and biases that can be associated with the mental construction of counterfactual alternatives to reality. As with other types of errors and biases in judgment, only a recognition of these errors and an understanding of the processes involved can improve our reasoning and decision-making.

REFERENCES

Baron, J. and Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, **54**, 569–579.

Boninger, D. S., Gleicher, F. and Strathman, A. (1994). Counterfactual thinking: From what might have been to what may be. *Journal of Personality and Social Psychology*, **67**, 297–307.

Davis, C. G., Lehman, D. R., Silver, R. C., Wortman, C. B. and Ellard, J. H. (1996). Self-blame following a traumatic event: The role of perceived avoidability. *Personality and Social Psychology Bulletin*, **22**, 557–567.

Gleicher, F., Kost, K. A., Baker, S. M., Strathman, A. J., Richman, S. A. and Sherman, S. J. (1990). The role of counterfactual thinking in judgments of affect. *Personality and Social Psychology Bulletin*, **16**, 285–295.

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., Ogawa, K. H., Delforge, A. and Early, D. (1989). Causal primacy and comparative fault: The effect of position in a causal chain in judgments of legal responsibility. *Personality and Social Psychology Bulletin*, **15**, 161–174.

Judd, C. M. and Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, **5**, 602–619.

Kahneman, D. and Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, **93**, 136–153.

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.

Markman, K. D., Gavanski, I., Sherman, S. J. and 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. and 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.

McGill, A. L. (1989). Context effects in judgments of causation. *Journal of Personality and Social Psychology*, **57**, 189–200.

McGill, A. L. and Klein, J. G. (1993). Contrastive and counterfactual thinking in causal judgment. *Journal of Personality and Social Psychology*, **64**, 897–905.

Miller, D. T. and 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. and 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. and 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.

Roese, N. J. (1994). The functional basis of counterfactual thinking. *Journal of Personality and Social Psychology*, **66**, 805–818.

Roese, N. J. and Olson, J. M. (1995). *What might have been: The social psychology of counterfactual thinking*. Mahwah, NJ: Lawrence Erlbaum.

Schwarz, N. and Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin and A. Tesser (Eds.), *The construction of social judgment* (pp. 213–245). Hillsdale, NJ: Lawrence Erlbaum.

Sears, D. O. (1983). The person-positivity bias. *Journal of Personality and Social Psychology*, **44**, 233–240.

Sherman, S. J. and McConnell, A. R. (1995). Dysfunctional implications of counterfactual thinking: When alternatives to reality fail us. In N. J. Roese and J. M. Olson (Eds.), *What might have been: The social psychology of counterfactual thinking* (pp. 199–231). Mahwah, NJ: Lawrence Erlbaum.

Sherman, S. J., Zehner, K. S., Johnson, J. and Hirt, E. R. (1983). Social explanation: The role of timing, set, and recall on subjective likelihood estimates. *Journal of Personality and Social Psychology*, **44**, 1127–1143.

Sherman, S. J., Beike, D. R. and Ryalls, K. R. (1995). *Inconsistencies in judgments of general versus specific cases*. Unpublished manuscript, Indiana University.

Spellman, B. A. (1995). *Crediting causality*. Unpublished manuscript, University of Texas.

Taylor, S. E., Lichtman, R. R. and Wood, J. V. (1984). Attributions, beliefs about control, and adjustment to breast cancer. *Journal of Personality and Social Psychology*, **46**, 489–502.

Thibaut, J. H. and Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley,

Wells, G. L. and Gavanski, I. (1989). Mental simulation of causality. *Journal of Personality and Social Psychology*, **56**, 161–169.