Journal of Experimental Psychology:
Learning, Memory, and Cognition
2000, Vol. 26, No. 5, 1318–1331

0278-7393/00/$5.00 DOI: 10.1037//0278-7393.26.5.1318

Counterfactual Thinking About Actions and Failures to Act

Ruth M. J. Byrne and Alice McEleney
University of Dublin, Trinity College

When people think counterfactually about how a situation could have turned out differently, they mentally undo events in regular ways (e.g., they focus on actions not failures to act). Four experiments examine the recent discovery that the focus on actions in the short term switches to inactions in the long term. The experiments show that this temporal switch occurs only for particular sorts of situations. Experiment 1 showed no temporal pattern to the agency effect when 112 participants judged emotional impact and frequency of "if-only" thoughts from both short- and long-term perspectives for an investment scenario. Experiment 2 showed no temporal pattern when 190 participants considered a college choice scenario with a good outcome. Experiment 3 showed no temporal pattern when 131 participants considered an investment scenario even when the situation for the actor and nonactor was bad from the outset. Experiment 4, with 113 participants, showed a focus on actions even when the investment loss was equal for both the actor and nonactor. The implications of the results are discussed in terms of what is explicitly available in the mental representation of actions and inactions.

Counterfactual thinking about how factual situations may have been different occurs often in everyday thinking (e.g., Kahneman & Miller, 1986). People spontaneously think about how an outcome could have turned out differently, and they mentally undo aspects of the events that led to it. Counterfactual thinking may serve a preparatory function in helping people to establish the causes of events and to learn from mistakes (e.g., Roese & Olson, 1995; Wells & Gavanski, 1989). It may also serve an affective function in contributing to emotions such as regret, guilt, and hope and attributions such as blame and responsibility (Kahneman & Tversky, 1982; Roese, 1994, 1997). This type of thinking plays a role in many aspects of cognition, including formulating counterexamples to deductive conclusions (e.g., Johnson-Laird & Byrne, 1991), deriving subgoals in problem solving (e.g., Ginsberg, 1986; Keane, 1997), and intending to improve in creative discovery (e.g., Hofstadter, 1985).

When people think about how an outcome could have been different, they mentally undo their representation of aspects of the factual situation in regular ways, perhaps corresponding to core categories of mental life, such as causal, temporal, spatial, and intentional categories (e.g., G. Miller & Johnson-Laird, 1976). For example, they mentally undo the first cause in a causal sequence rather than subsequent causes (e.g., Wells, Taylor, & Turtle, 1987), but they mentally undo the most recent event in a noncausal, temporal sequence rather than earlier events (e.g., Byrne, Segura, Culhane, Tasso, & Berrocal, 2000; D. T. Miller & Gunesagaram, 1990; Spellman, 1997). They mentally undo events under a person's voluntary control rather than events outside their control (e.g., Girotto, Legrenzi, & Rizzo, 1991; McCloy & Byrne, in press). Some research has indicated that they tend to judge that people will feel worse about their actions rather than their failures to act, perhaps because mentally undoing an event amplifies its emotional impact (Kahneman & Tversky, 1982). Our aim was to examine this last phenomenon: the tendency to judge that people will feel worse about matters of commission rather than omission. This *agency effect,* as we call it, seems a pervasive and potent feature of everyday experience: It has been established in both laboratory studies and everyday naturalistic situations (e.g., Gilovich & Medvec, 1995a), and it may yield important insights into the nature of the psychological mechanisms underlying cognitive and emotional processes in counterfactual thinking.

The Agency Effect

Kahneman and Tversky (1982, p. 173) showed that reasoners judge that actors will regret their actions more than nonactors will regret their inactions. Ninety-two percent of participants in their experiment believed that the individual who acted in the following scenario, Mr. George, would feel greater regret:

> Mr. Paul owns shares in company A. During the past year he considered switching to stock in company B, but he decided against it. He now finds out that he would have been better off by $1,200 if he had switched to the stock of company B. Mr. George owned shares in company B. During the past year he switched to stock in company A.

Ruth M. J. Byrne and Alice McEleney, Department of Psychology, University of Dublin, Trinity College, Dublin, Ireland.

This research was supported in part by grants from the Dublin University Arts and Social Sciences Benefactions Fund and the Provost's Academic Development Fund, to Ruth Byrne. Alice McEleney is supported by an Enterprise Ireland Ph.D. fellowship, a Dublin University Foundation Scholarship, and a Government of Ireland Scholarship for the Humanities and Social Sciences. The results of some of the experiments were reported at the 19th annual conference of the Cognitive Science Society at Stanford University, 1997, and the International Workshop on Reasoning: Logical and Psychological Perspectives, Padua University, 1999.

We thank Mark Keane, Vittorio Girotto, Phil Johnson-Laird, and Rachel McCloy for their helpful comments on this research. We are grateful to Bobbie Spellman and an anonymous reviewer for their suggestions leading to the third and fourth experiments and to Clare Walsh for her help with those experiments.

Correspondence concerning this article should be addressed to Ruth M. J. Byrne, Department of Psychology, University of Dublin, Trinity College, Dublin 2, Ireland. Electronic mail may be sent to rmbyrne@tcd.ie.

> He now finds out that he would have been better off by $1,200 if he had kept his stock in company B. Who feels greater regret?

Of course, the objective situation for each individual is entirely the same, but people judge that their subjective situation is quite different. Even when the order of the actions and inactions is varied in the scenarios, people judge the action to be regretted rather than the inaction; that is, people do not simply focus on the more recently mentioned event (Gleicher et al., 1990). People readily judge that an actor regrets actions and a nonactor regrets inactions when these judgments are elicited separately, but when a comparative judgment is required, they judge that it is the actor who experiences more regret, at least when the actions and inactions play a causal role in bringing about the negative outcome (N'Gbala & Branscombe, 1997).

Although people judge that actions are more regretted than failures to act in the short term, they judge that failures to act are more regretted than actions in the long term (Gilovich & Medvec, 1994, 1995a). Gilovich and Medvec (1994, p. 360) demonstrated this reversal of the agency effect, from a focus on actions to a focus on inactions, with the following scenario:

> Dave and Jim do not know each other, but both are enrolled at the same elite East Coast university. Both are only moderately satisfied where they are and both are considering transferring to another prestigious school. Each agonizes over the decision, going back and forth between thinking he is going to stay and thinking he will leave. They ultimately make different decisions: Dave opts to stay where he is, and Jim decides to transfer. Suppose their decisions turn out badly for both of them: Dave still doesn't like it where he is and wishes he had transferred, and Jim doesn't like his new environment and wishes he had stayed.

In their experiment, 76% of participants who were asked the question "Who do you think would regret his decision more upon learning that it was a mistake?" thought that Jim, the individual who had acted, would feel more regret. However, 63% of the participants who were asked the question "Who do you think would regret his decision more in the long run?" thought that Dave, the individual who failed to act, would feel more regret.

Of course, judgments about the emotional experience of a fictional character in a scenario are very different from genuinely experienced emotions in a similar real-life situation. Scenario-based emotion judgments may be based on inferential processes and influenced by memory, perspective, suppositions, and so on. Nonetheless, regrets for real-life actions and inactions appear to increase the distress experienced in grief (e.g., Davis, Lehman, Wortman, Silver, & Thompson, 1995; see also Dunning & Parpal, 1989). Moreover, people's real-life regrets for events recalled from their own past lives do seem to follow a temporal pattern (Gilovich & Medvec, 1994). When people are asked to recall events they regret from their past, their regrets in the past week are evenly split between actions and failures to act, but their regrets looking back over their lives to that point are mainly regrets for what they failed to do. They tend to say that they regret most failing to spend time with their family and friends, failing to avail of educational opportunities, failing to pursue hobbies, and so on (Gilovich & Medvec, 1994).

The temporal pattern to the experience of regret may result from a variety of variables. Our aim in this article was to examine some of the cognitive factors that mediate the agency effect. Gilovich and Medvec (1994) proposed that cognitive variables that may play a role include the following: (a) People remember regretted inactions better than actions (Gilovich & Medvec, 1995a; Savitsky, Medvec, & Gilovich, 1997), and (b) the consequences of regretted inactions remain "open," that is, factually unknown. The source of regret for actions may be their factual consequences, whereas the source of regret for inactions may be their counterfactual consequences, which are bounded only by the limits of the imagination of what might have been (Gilovich & Medvec, 1995a). We examine this latter variable, the role of imagined consequences, in the four experiments we report here.

Of course, other variables, including motivational and emotional variables, may also be important. Gilovich and Medvec (1994) proposed that seven variables are crucial in the agency effect: the two cognitive variables outlined earlier and five motivational variables that might reduce regret for action with time: (a) People try to compensate for bad outcomes from actions by trying to set them right; (b) they try mentally to look for "silver linings" to bad outcomes from actions; (c) They try mentally to reduce dissonance arising from actions, and that might increase regret for inaction with time; (d) people find inactions inexplicable in the light of retrospective confidence about the success of action; and (e) people find it difficult to access the variables that inhibited action. Emotional variables may include the possibility that short-term regrets may be "hot," whereas long-term regrets are wistful (Kahneman, 1995). In fact, people judge that regret for actions is accompanied by anger, whereas regret for inactions is accompanied by nostalgia or misery (Gilovich, Medvec, & Kahneman, 1998). We return to a consideration of such variables, after we have considered the crucial cognitive processes underlying the agency effect.

Real and Imagined Consequences

We suggest that the tendency to judge greater emotional impact for actions may result from the way in which actions are mentally represented. People may construct mental representations that make some information explicit and leave other information implicit, as suggested by studies of text comprehension (e.g., Garnham, 1985) and reasoning (e.g., Johnson-Laird & Byrne, 1991). Our first proposal is that actions may be represented more explicitly—that is, by keeping in mind more information—than inactions (e.g., Byrne & McEleney, 1997; Kahneman & Miller, 1986). Actions may be represented by keeping in mind the preaction state (e.g., Jim in College A), and by also keeping in mind the postaction state (e.g., Jim in College B). In contrast, inactions can be represented by keeping in mind just a single state (e.g., Dave in College A) because there has been no change in that state. This cognitive difference may be the underlying root of the perception that actions are a departure from the status quo (Kahneman & Miller, 1986). Because of the greater amount of information kept in mind about actions, they may be more readily available for mutation in the generation of a counterfactual alternative (e.g., the postaction state can be replaced with the past, now counterfactual, preaction state to generate the thought, "If only Jim had stayed in College A"). Our first key proposal is that actions are more readily mutated because they are represented explicitly.

Why then is there a reversal to a focus on inactions from a long-term perspective? People can revise their mental representations of events to make implicit information explicit if need be. They carry out such a revision when they realize that they have represented insufficient information for the task at hand. When they are given the task of judging who will feel worse about their situation from a long-term perspective, they must consider the real and imagined consequences of the events, because over time, it is an event's consequences rather than the event itself that may continue to have an impact (e.g., Shafir & Tversky, 1992). For the actor, the consequences of each college choice are known (e.g., unhappiness for Jim in College A and even more unhappiness in College B), but for the nonactor the consequences of one college choice only are known (e.g., unhappiness for Dave in College A). When people revise their mental representation of the events to imagine the consequences of the other college choice (Dave in College B), there are at least two possibilities (Dave could have been happy or unhappy in College B), and one of them (Dave's imagined happiness) is vastly better than the real consequences (Dave's current unhappiness). When reasoners revise their mental representations, they discover that the consequences of the mentally undone inaction are possibly good. In the counterfactual situation, Dave might have been happy, whereas in the factual situation, Dave is unhappy. Because the imagined consequences are possibly good for Dave, they conclude that he feels worse (he could have been happy had he acted differently) than Jim feels (he would have been unhappy even if he had acted differently); and because the consequences are possibly good for Dave, reasoners focus on his inaction (more often than Jim's action) when they generate a counterfactual alternative, say, in a sentence-completion task (e.g., "If only Dave had transferred, he might have been happy"). The college choice scenario provides unmatched information about the action and the inaction: The factual consequences of the action and the counterfactual consequences of the mentally undone action are both known and the same (Jim was unhappy in Colleges A and B). However, the factual consequences of the inaction are known (Dave was unhappy in College A), but the counterfactual consequences of the mentally undone inaction are unknown and possibly better (Dave might have been happy in College B). Therefore, our second key proposal is that inactions are more readily mutated from the long-term perspective only when their counterfactual consequences are possibly better than their real consequences (and the action's counterfactual consequences are not).

The college scenario is crucially different from the investment scenario in this respect (Kahneman & Tversky, 1982; Landman, 1987). In the investment scenario, the factual and counterfactual consequences of the action and inaction are both known and similar (e.g., Kahneman, 1995). For the actor, the factual and counterfactual consequences for each investment choice are known (he is in Company B to start with and he transfers to Company A; he loses money in Company A, which he would not have lost in Company B); likewise, for the nonactor, the factual and counterfactual outcomes for each investment choice are also known (he is in Company A to start with and he stays in Company A; he loses money in Company A, which he would not have lost in Company B). Our central claim is that the temporal pattern to the agency effect is confined to a special set of circumstances: situations in which the imagined consequences from the mentally undone action are the same as the real consequences from the action, but the imagined consequences from the mentally undone inaction are possibly better than the real consequences from the inaction. From a long-term perspective, people consider the consequences of actions and inactions, and the discrepancy between them in the college scenario becomes apparent. The real consequences of the inaction may seem worse and, so, are judged to be regretted more than the real consequences of the action, compared with the imaginary consequences. The actor can be judged to consider the possibility "If I had stayed in College A, I would have been unhappy," because it is factually known that when he was in College A in the past, he was unhappy. In contrast the nonactor can be judged to consider the possibility "If I had switched to College B, I might have been happy or I might have been unhappy." The counterfactual possibility of happiness makes the real outcome seem worse for the nonactor than the actor. Of course, in the short term, people concentrate on what they have explicitly represented about the actions and inactions rather than their real or imaginary consequences and so, the discrepancy is not apparent. In all other situations, all things being equal, actions should be more mutable regardless of temporal perspective. If our account is correct, then the temporal reversal should not occur for the investment scenario.

However, perhaps the "open" nature of the college scenario—that is, the fact that the imagined consequences of the mentally undone inaction are unknown—is more similar to real-life situations (Gilovich & Medvec, 1995b)? After all, regret for inactions in the long term has been demonstrated for people's regrets from their own lives (Gilovich & Medvec, 1995a). In fact, experimental evidence indicates that people characterize regrets for actions and inactions from their own life more often as "closed" rather than as open (Savitsky et al., 1997, Study 3). Savitsky and his colleagues asked participants to describe their three biggest regrets of action and three of inaction and to rate on a scale from *closed book* to *unfinished business* the extent to which each regret was closed- or open-ended. Participants considered 40% of their regrets to be open and 59% to be closed. Hence, a closed scenario, such as the investment one, has as much in common with natural regrets for actions and inactions as an open scenario, such as the college one.

Overview of the Experiments

Here we report the results of four experiments in which we examined our proposals that actions are more often mutated than inactions except when the imagined consequences of the mentally undone inaction are possibly better than the real consequences of the inaction (and the imagined consequences of the mentally undone action are the same as the real consequences of the action). In Experiment 1, we tested the novel prediction derived from our account that there should be no temporal reversal to the agency effect for the investment scenario. In the investment scenario, the imagined consequences of the mentally undone inaction are known to be better than the real consequences of the inaction, and exactly the same is true for the action: The amount and nature of the information about the action and inaction are matched. In Experiment 2, we tested the novel prediction that there should be no temporal reversal to the agency effect for good outcomes even for the college scenario. In a college scenario with a good outcome, the imagined consequences of the mentally undone inaction are no better than the real consequences of the inaction. The third exper-

iment pits our explanation of the temporal reversal against an alternative based on differences in the initial states of the protagonists (which are neutral in the investment scenario but bad in the college scenario). In Experiment 4, we tested our explanation against an alternative that was based on differences in the subjective outcomes for the actor and the nonactor. The experiments were guided by our two central proposals: (a) People mentally undo actions because actions are represented more explicitly in their minds, and (b) they switch to mentally undoing inactions only in special circumstances (when the imagined consequences of the mentally undone inaction are better than the real consequences of the inaction, whereas the imagined and real consequences for the action are the same).

Experiment 1: Temporal Perspectives on Bad Investments

The crucial differences between the college scenario and the investment one are whether the counterfactual consequences of the action and the inaction are both known and whether both differ from the real consequences in similar ways. Given these differences, our aim in Experiment 1 was to examine whether people mentally undo actions from a short-term perspective and inactions from a long-term perspective for the investment scenario in which the real and imagined consequences from the action and inaction are matched or whether this temporal pattern is confined only to scenarios of the college structure, in which the real and imagined consequences from the action and inaction are not matched. We gave participants investment scenarios from a short-term or a long-term perspective, and we asked them to judge the emotional impact of the outcome on the protagonists.

In addition, we asked the participants who would think "if only" most and how they would complete this thought. The agency effect has been established primarily through judgments of emotion, with the assumption that emotions such as regret are cognitively determined emotions that are based on counterfactual mutability (e.g., Gilovich & Medvec, 1995a; Kahneman & Miller, 1986). The exact nature of the link between emotional amplification and counterfactual mutability has only recently begun to be studied for the agency effect (e.g., N'Gbala & Branscombe, 1997). In our mutation task, the frequency of participants' judgments that the actor or the nonactor would think "if only" allowed us to probe subtly and indirectly which person's actions or nonactions the participants considered to be most mutable.

Method

Materials and design. We constructed a short-term and a long-term version of a scenario based on Kahneman and Tversky's (1982) investment scenario. The short-term scenario was as follows (the information for the long-term scenario is provided in parentheses; it was given to participants in addition to the information for the short-term scenario):

> Paul saved hard and had £2,000 to invest. He bought shares in Company A. During his first year as a shareholder in Company A there was a once-off offer to switch to shares in Company B, but he decided against it. At the end of the first year he found out that he would have been better off by £5,000 if he had taken the offer and switched to shares in Company B. (After 10 years, he found out that his shares did not make up the lost ground in the meantime, and he would still have been better off by £5,000 if he had switched to shares in Company B.)
>
> Joe also saved hard and had £2,000 to invest. He bought shares in Company B. During his first year as a shareholder in Company B there was a once-off offer to switch to shares in Company A, and he decided to do so. At the end of the first year he too found out that he would have been better off by £5,000 if he had turned down the offer and kept his shares in Company B. (After 10 years, he found out that his shares did not make up the lost ground in the meantime, and he would still have been better off by £5,000 if he had kept his shares in Company B.)

The participants were asked three questions. In the short-term condition, they were asked the following questions:

> At the end of the first year of investment:
> 1. Who do you think would imagine "if only . . . " most often at the end of the first year of investment?
> Please tick one: Paul ____ Joe ____
> 2. How do you think he completed this thought?
> 3. Who do you think would feel worse about his decision at the end of the first year of investment?
> Please tick one: Paul ____ Joe ____

The participants given the long-term scenario were asked the same questions but "at the end of the first year of investment" was replaced by "after 10 years of investment." They were asked the three questions in the fixed order given here. We gave the short-term scenario to one group of participants and the long-term scenario to a second group.

Participants and procedure. The 120 participants were undergraduate students from different departments in the University of Dublin, Trinity College, Dublin, Ireland. The 98 women and 22 men, aged 17 to 54 years (M = 20 years), participated in the experiment voluntarily and were assigned at random to two groups. Eight women were eliminated prior to any data analysis because they failed to complete all three questions. The remaining 112 participants completed all three questions for either a short-term or a long-term scenario (n = 56 in each). The participants were tested in several groups and were presented with a booklet that contained the instructions, the scenario and questions, and a debriefing paragraph, each on separate pages. The participants worked at their own pace and provided their answers in the spaces on the sheet.

Results and Discussion

The results showed that people judged greater emotional impact and greater frequency of counterfactual thoughts for actions in the short term and in the long term, when the counterfactual consequences of the actions and inactions were both known and the same. From the short-term perspective, 87.5% of the participants judged that the actor, Joe, would feel worse, whereas 12.5% judged that the nonactor, Paul, would feel worse, and the difference was reliable (binomial $z = 5.61$, $p < .01$). Likewise from the long-term perspective, 86% of the participants judged that the actor, Joe, would feel worse whereas 14% judged that the nonactor Paul, would feel worse and the difference was reliable (binomial $z = 5.35$, $p < .01$). There was no difference in the short-term and long-term conditions in the judgments that the actor would feel worse (87.5% vs. 86%, $\chi^2(1, N = 112) = 0.07$, $p > .10$, as Table 1 shows (and the comparison had sufficient power at the 80% level to detect a difference of 18% or more).

Exactly the same pattern emerged for participants' judgments about who would think "if only" most often. From the short-term perspective, people judged that the actor Joe (71%) would think "if only" more often than the nonactor Paul (29%), and the difference was reliable (binomial $z = 3.21$, $p < .01$). Likewise, from the

Table 1
Percentage of Judgments of Who Feels Worse and Who Thinks "If Only" Most Often for Actions and Inactions in the Short and Long Terms for Investment Scenarios With Known Outcomes

Question	Short term	Long term
Who feels worse?		
Action	87.5	86
Inaction	12.5	14
Who thinks "if only" most often?		
Action	71.0	64
Inaction	29.0	36
n	56	56

long-term perspective, people judged that the actor Joe (64%) would think "if only" more often than the nonactor Paul (36%), and the difference was reliable (binomial $z = 2.14, p < .01$). There was no difference in the short-term and long-term conditions in the judgments that the actor would think "if only" most (71% vs. 64%), $\chi^2(1, N = 112) = 0.65, p > .10$, as Table 1 shows (and the comparison had sufficient power at the 80% level to detect a difference of 25% or more).

People completed the counterfactual thought by undoing the individual's action or inaction (depending on whether they focused on the actor or the nonactor) in 80% of cases for the short-term scenario, and the remaining 20% of mutations were general ones (e.g., "I wish I'd taken a chance" or "I wish I had thought harder"). None of the undoings focused on the consequences (e.g., "I wish I hadn't lost the money"). Likewise, from the long-term perspective, people completed the thought by undoing the individual's action or inaction in 70% of cases, and the remaining 30% of mutations were general (e.g., "I wish I hadn't been greedy" or "I wish I had been more patient"). Once again, none of the undoings focused on the consequences.

The results show that people do not mentally undo inactions from a long-term perspective when their counterfactual consequences are known and the same for both protagonists. Actions are considered to have greater emotional impact and to generate greater frequency of counterfactual thoughts than inactions from both a long-term and a short-term perspective, when the counterfactual consequences are known and the same for both the action and the inaction. The results also show that judgments of who feels worse and frequency of "if-only" thoughts follow the same pattern as each other (N'Gbala & Branscombe, 1997). An increase in counterfactual emotion has been assumed to correspond to an increase in the mutability of an event (e.g., Kahneman & Miller, 1986; Roese & Olson, 1995), and our results are consistent with this suggestion (of course, the within-participant design may have amplified the internal consistency of participants' judgments in this respect).

The results support our proposal that the focus on inactions in the long term for the college scenario may depend on the relative amount and nature of information available about the counterfactual consequences of the action and inaction. It remains unknown whether Dave, the nonactor, would have been happier if he had moved to another college (Gilovich & Medvec, 1995b; Kahneman, 1995). The counterfactual consequences of some sorts of inactions, such as whether a move to another college would have made a person happier, are inherently unknowable (one individual's life, as psychologists have long noted, has no control comparison); but the counterfactual consequences of other sorts of inactions, such as whether an investment in another company would have made a profitable return, are entirely knowable. Inactions may have more emotional impact than actions from a long-term perspective only for the former sort of inactions, ones for which counterfactual consequences are unknown and possibly better than the real consequences.

One alternative possible explanation is that people may look for silver linings from experiences that turn out badly (Gilovich & Medvec, 1994), and the actor in the college scenario may appear to gain more from his experience than the actor in the investment scenario. The actor in the college scenario gains experience of two colleges, and even though both result in unhappiness, the additional experience may be viewed as a bonus in the long term. The nonactor gains experience of just one college and so may be judged to have more to regret. Likewise, the actor in the investment scenario gains experience as a shareholder in two companies, and the nonactor gains experience of just one company, but perhaps the additional experience in the case for the actor may not confer much compensatory experience. However, this alternative possible explanation is not consistent with the results of the experiment. If the reversal to a focus on inactions from the long-term perspective in the college scenario resulted mainly from the availability of silver linings, then the investment scenario should lead to an elimination of any agency effect, with the focus being equally on actions and inactions, from the long-term perspective because neither the actor nor the nonactor gained more than one another in this regard. Instead, there was a robust focus on actions for the investment scenario.

Experiment 1 showed that there is no temporal pattern to the judgments of emotional impact and frequency of counterfactual thoughts when the consequences of actions and inactions are both known and similar. The temporal pattern to the agency effect—the reversal from a focus on actions from a short-term perspective to inactions from a long-term perspective—is confined to certain sorts of scenarios, that is, scenarios in which the counterfactual consequences for the nonactor are unknown and possibly better than the real consequences.

Experiment 2: College Choices With Good Outcomes

The results of the first experiment supported our suggestion that the temporal reversal to inactions occurs only when the counterfactual consequences for the nonactor are unknown and possibly better than the real consequences (Dave is unhappy in College A; he might have been happy in College B) in contrast to the counterfactual consequences for the actor that are known and the same (Jim is unhappy in Colleges A and B). In some situations, the counterfactual consequences for the nonactor may be unknown but possibly worse than the real consequences (e.g., Dave is happy in College A; he might have been unhappy in College B). By our account, no reversal should occur in such situations, even though the counterfactual consequences are unknown. The inaction will not seem to have led to a worse outcome than the action (compared with their counterfactual consequences), and so it will not be focused on more. Recall that for the bad-outcome scenario, the unhappy actor could be judged to consider "If I had stayed in

College A, I would have been unhappy" (because he was unhappy when he was there in the past), whereas the unhappy nonactor could be judged to consider "If I had switched to College B, I might have been happy or I might have been unhappy." Our suggestion is that the counterfactual possibility of happiness makes the real outcome (of unhappiness) seem worse for the nonactor than the actor. For a good-outcome scenario the happy actor could be judged to consider "If I had stayed in College A, I would have been unhappy," whereas the happy nonactor could be judged to consider "If I had switched to College B, I might have been happy or I might have been unhappy." The counterfactual possibility of happiness does not make the real outcome (of happiness) seem worse for the nonactor than the actor. To test this proposal, we examined college scenarios with good outcomes as well as bad outcomes, from the short- and long-term perspective.

A focus on inactions, rather than actions, from a long-term perspective has been observed experimentally only for the college scenario, and in fact only a college scenario with a bad outcome has been investigated (Gilovich & Medvec, 1994, 1995a). A focus on actions in the short-term has been observed for scenarios with either good or bad outcomes (Landman, 1987). For example, people judge that actors will feel better about actions that lead to good outcomes than nonactors will feel about inactions. Landman demonstrated the emotional impact of actions for good outcomes with a variety of scenarios, including the following:

> Paul enrolled in Section 1 of Biology 101; his roommate enrolled in Section 2. At the beginning of the term, Paul considered switching to Section 2, but decided against it. The term is over and Paul just learned that he got an A in the course. His roommate got a B. George and his roommate enrolled in Section 2 of Biology 101. At the beginning of the term, George switched to Section 1. The term is over and George just learned that he got an A in the course. His roommate got a B in Section 2. Who felt better about his section choice, Paul or George?

Sixty-seven percent of participants judged that George, the individual who acted, would feel better about his section choice than Paul, the individual who did not act (Landman, 1987). Is there a temporal pattern to the agency effect for good outcomes? The aim of our second experiment was to address this question: We examined whether the temporal reversal of the agency effect for the college scenario with a bad outcome would also be observed for a college scenario with a good outcome. We also made sure that the scenarios specified identical long-term outcomes for the actor and the nonactor in the bad-outcome conditions to control for any ambiguities in this feature in previous research.

Method

Materials and design. We constructed four scenarios on the basis of fictional characters John and Paul and their college choices: a short-term, bad-outcome version; a long-term, bad-outcome version; a short-term, good-outcome version; and a long-term, good-outcome version. We constructed the four scenarios based on Gilovich and Medvec's (1994) college scenario, which we altered in several ways to clarify the finality of the decision and especially the long-lasting impact of the outcome:

> John and Paul do not known each other, but both are enrolled at the same elite university. Both are only moderately satisfied where they are and both are considering transferring to another prestigious university. They are both told they must make their final decision before the end of the year. Each agonizes over the decision, going back and forth between thinking he is going to stay and thinking he will leave. They ultimately make different decisions: John opts to stay where he is, and Paul decides to transfer.
>
> Suppose their decisions turn out badly for both of them: At the end of the year, John is even more unhappy where he is and wishes he had transferred, and Paul is even more unhappy at his new college and wishes he had stayed where he was. However it is too late for either of them to reverse his decision.

The good-outcome version continued with a different second paragraph as follows:

> Suppose their decisions turn out well for both of them: At the end of the year, John is happier where he is and is glad he stayed where he was, and Paul is happier at his new college and is glad he transferred. Who do you think would feel better about his decision on learning that it was the right one?

The participants' task was to tick the name of one of the individuals, John or Paul, in response to a question at the end of the scenario. Participants in the short-term conditions were asked the following question:

> Who do you think would regret his decision more [feel better about his decision] on learning it was a mistake [it was the right one]? John ____ Paul ____

Participants in the long-term conditions were asked the following question, for the bad-outcome version:

> As a result they both drop out of college and neither of them ever secures a good job. Who do you think would regret his decision more on looking back on it ten years later? John ____ Paul ____

and for the good-outcome version:

> They both do very well at college and secure good jobs after graduating. Who do you think would feel better about his decision looking back on it ten years later? John ____ Paul ____

Participants were given one of four versions of the scenarios at random.

Participants and procedure. The 190 participants were undergraduate students from the University of Dublin, Trinity College. The 159 women and 31 men, aged 16 to 54 years (mean age = 21), participated in the experiment voluntarily. They were assigned at random to the short-term bad-outcome condition (n = 46), long-term bad-outcome condition (n = 46), short-term good-outcome condition (n = 47), or the long-term good-outcome condition (n = 51). They were tested in several groups. Participants were presented with a 2-page booklet that contained the scenario and question on the first page, followed by a debriefing paragraph on the second page. They were instructed to take their time reading the story and to answer the question that followed. The participants worked at their own pace and ticked their answers in the spaces provided on the sheet.

Results and Discussion

The results showed a temporal reversal for the agency effect for bad outcomes: Participants mentally undid actions in the short-term, bad-outcome scenario but inactions in the long-term, bad-outcome scenario, replicating Gilovich and Medvec's (1994) results. However, the results showed no temporal reversal for the agency effect for good outcomes: Participants mentally undid actions in both the short-term, good-outcome and the long-term, good-outcome scenarios.

As Table 2 shows, a majority of the participants (61%) who received the short-term bad-outcome scenario indicated that the actor Paul would feel more regret than the nonactor John (39%; binomial $z = 1.47$, $p < .10$). In contrast, most of the participants who received the long-term, bad-outcome version indicated that the nonactor John would feel more regret (76%) than the actor Paul (24%); binomial $z = 3.54$, $p < .01$). More participants judged that the actor would feel more regret in the short-term scenario (61%) than in the long-term scenario (24%), and this difference was reliable $\chi^2(1, N = 92) = 12.86$, $p < .001$. Thus, the experiment replicated Gilovich and Medvec's (1994) finding of a reversal from mentally undoing actions to mentally undoing inactions over time for bad outcomes.

As Table 2 also shows, this reversal did not occur for good outcomes. Most participants who received the short-term, good-outcome version indicated that the actor Paul would feel better about his decision (70%) rather than the nonactor John (30%; binomial $z = 2.80$, $p < .01$). Likewise, most participants who received the long-term, good-outcome version indicated that the actor Paul would feel better about his decision (75%) rather than the nonactor John (25%; binomial $z = 3.50$, $p < .01$). There was no difference between judgments that the actor Paul would feel better in the short-term scenario (70%) than in the long-term scenario (75%), $\chi^2(1, N = 98) = 0.23$, $p > .10$ (and the comparison had sufficient power at the 80% level to detect a difference of 25%).

The results showed that for bad outcomes, there was a temporal pattern to the agency effect for the college scenario with its unknown and possibly better counterfactual consequences for the inaction: Participants mentally undid actions from the short-term perspective and inactions from the long-term perspective. But for good outcomes, there was no temporal pattern to the agency effect, even for the college scenario, when the counterfactual consequences for the inaction were unknown but possibly worse: Participants mentally undid actions from both the short-term and long-term perspectives.

Perhaps the wording of the long-term scenarios led participants to consider that the consequences of the bad outcome were entirely known, whereas the consequences of the good outcome were less known? The long-term, bad-outcome scenario contained the information, "Neither of them ever secures a good job," whereas the long-term, good-outcome scenario contained the information, "They both . . . secure good jobs after graduating;" it remains unknown whether they lost their jobs in the intervening 10 years, stayed in them, or got better ones. But this possible difference in the scenarios is unlikely to explain the results of the experiment. In fact, it erroneously predicts a temporal reversal for good outcomes: If the counterfactual consequences for the inaction were even more unknown for the good-outcome scenario than the bad-outcome scenario, then there should have been an even greater reversal of the agency effect for the good-outcome scenario.

Table 2
Percentages of Judgments of Greater Emotion for Actions and Inactions in the Short and Long Terms for College Scenarios With Good and Bad Outcomes

Outcome	Short term	Long term
Bad outcome		
Action	61	24
Inaction	39	76
n	46	46
Good outcome		
Action	70	75
Inaction	30	25
n	47	51

Perhaps the wording of the scenarios differed in how explicit they made the counterfactual alternatives? In the good-outcome versions, the wording alluded only to the factual situation:

> John is happier where he is and is glad he stayed where he was, and Paul is happier at his new college and is glad he transferred.

whereas for the bad-outcome scenarios, the wording provided an explicit counterfactual scenario:

> John is even more unhappy where he is and wishes he had transferred, and Paul is even more unhappy at his new college and wishes he had stayed where he was.

An explicit counterfactual scenario enhances the emphasis on actions, at least for good outcomes (Gleicher et al., 1990). We ruled out any alternative explanations of our results on the basis of differences in wording in the two scenarios in a further experiment. We gave 257 participants good-outcome scenarios from short-term and long-term perspectives, similar to the ones used in this experiment, and we compared them with good-outcome, short- and long-term scenarios that contained explicit counterfactual wording:

> John is happier where he is and is glad he had not transferred, and Paul is happier at his new college and is glad he had not stayed where he was.

The results showed that the explicit counterfactual wording yielded the same pattern of results as the scenario with no explicit counterfactual wording: Participants mentally undid the action more than the inaction in the short term (explicit counterfactual, 84% vs. 16%; no explicit counterfactual, 68% vs. 32%) and in the long term (explicit counterfactual, 88% vs. 12%; no explicit counterfactual, 73% vs. 27%). If anything, the explicit counterfactual accentuated the results rather than diminished them (see Byrne & Gillanders, 1999).

The experiment showed that the temporal reversal for the agency effect was confined not only to situations with the structure of the college scenario but more specifically only to such situations when they lead to a bad outcome. Our account proposes that people undo actions because they are mentally represented in an explicit way; they switch to undoing inactions when they consider the consequences of the events and discover that, whereas the real consequences of the action and the counterfactual consequences of the mentally undone action are known and the same, the real consequences of the inaction are known but the counterfactual consequences of the mentally undone inaction are unknown and are possibly better than the real consequences. If the unknown counterfactual consequences of the mentally undone inaction are unknown but are possibly worse than the real consequences, as they are in the good-outcome scenarios, there is no switch to a focus on inactions. The results support our suggestion that the

reversal depends on the counterfactual consequences of the inaction being unknown and possibly better.

Alternative explanations of the failure to find a temporal pattern to the agency effect in the investment scenario hinge on identifying other differences between the investment and college scenarios as critical to the effect, and we turn now to examine one such alternative explanation.

Experiment 3: Temporal Perspectives on Investments That Go Bad From the Outset

The temporal pattern to the agency effect occurs only for scenarios in which the counterfactual consequences for the inaction are unknown and possibly better than the real consequences. There is no temporal pattern to the agency effect for scenarios in which the counterfactual consequences of the inaction are known (the investment scenarios) or in which the counterfactual consequences of the inaction are unknown but are possibly worse than the real consequences (the good-outcome college scenarios). Are there any alternative explanations for the observation of a temporal reversal for the college scenario but not for the investment scenario? One possible difference between the two sorts of scenarios is that the college scenario is bad from the outset: Both individuals are unhappy in their college. In contrast, the investment scenario is neutral at the outset: No information is given about whether the individuals are happy with their investments. Perhaps participants' judgments about emotional impact depend in part on their view of whether the protagonists were in an unhappy situation to start with? Someone who is unhappy and does nothing about it may be judged to feel worse about their inaction over time than someone who is unhappy and tries to change their state, even if they do not succeed (e.g., Zeelenberg, 1996). The increased emotional impact of inactions over time may depend on the unhappy starting point of the protagonists' situations. In this experiment, we tested this alternative explanation.

We constructed an investment scenario that had a bad outset: Both individuals' share values have fallen. Instead of the following neutral outset in the original investment scenario:

> During his first year as a shareholder in Company A/B there was a once-off offer to switch to shares in Company B/A, but he decided against it/and he decided to do so

we provided a bad outset:

> During his first year as a shareholder in Company A/B the value of his shares fell considerably until they were worth only £1,000. Around that time there was a once-off offer to switch to shares in Company B/A, but he decided against it/and he decided to do so.

If the bad outset leads people to believe with hindsight (i.e., in the longer term) that the nonactor should have tried to do something to change his bad situation, then there should be a temporal pattern to the agency effect for this bad-outset scenario: Participants should switch from a focus on actions in the short term to a focus on inactions in the long term. The experiment was otherwise similar to the first one and compared who would feel worse and who would think "if only" most from the short-term and long-term perspectives.

Method

Materials and design. We constructed short-term and long-term versions of the scenario used in Experiment 1 that were similar to the scenarios described earlier except that they had a bad outset:

> Paul saved hard and had £2,000 to invest. He bought shares in Company A. During his first year as a shareholder in Company A the value of his shares fell considerably until they were worth only £1,000. Around that time there was a once-off offer to switch to shares in Company B, but he decided against it. At the end of the first year he found out that he would have been better off by £5,000 if he had taken the offer and switched to shares in Company B. [After 10 years, he found out that his shares did not make up the lost ground in the meantime, and he would still have been better off by £5,000 if he had switched to shares in Company B.]

> Joe also saved hard and had £2,000 to invest. He bought shares in Company B. During his first year as a shareholder in Company B the value of his shares fell considerably until they were worth only £1,000. Around that time there was a once-off offer to switch to shares in Company A, and he decided to do so. At the end of the first year he too found out that he would have been better off by £5,000 if he had turned down the offer and kept his shares in Company B. [After 10 years, he found out that his shares did not make up the lost ground in the meantime, and he would still have been better off by £5,000 if he had kept his shares in Company B.]

The participants were asked three questions about who would think if only most often, how the thought would be completed, and who would feel worse. We gave the short-term scenario to one group of participants and the long-term scenario to a second group.

Participants and procedure. The 131 participants were undergraduate students from different departments at the University of Dublin, Trinity College. The 65 women and 66 men, aged 18 to 50 years (M age = 24 years), participated in the experiment voluntarily and were assigned at random to one of two groups. Thirty participants were eliminated before any data analysis because they failed to complete all three questions. The remaining 101 participants completed all three questions for either the short-term (n = 50) or long-term scenario (n = 51). They were tested in several groups and presented with a booklet that contained the instructions, the scenario and questions, and a debriefing paragraph, each on separate pages. The participants worked at their own pace and provided their answers in the spaces on the sheet.

Results and Discussion

The results showed that people judged greater emotional impact and greater frequency of counterfactual thoughts for actions in the short term and in the long term, even when the starting point was bad. From the short-term perspective, people judged that the actor Joe (70%) would feel worse than the nonactor, Paul (30%), and the difference was reliable (binomial $z = 2.69$, $p < .004$). Likewise, from the long-term perspective, people judged that the actor Joe (76%) would feel worse than the nonactor Paul (24%), and the difference was again reliable (binomial $z = 3.64$, $p < .0002$). There was no difference in the short-term and long-term conditions in the judgments that the actor would feel worse (70% vs. 76%), $\chi^2(1, N = 101) = 0.46$, $p > .54$, as Table 3 shows (and the comparison had sufficient power at the 80% level to detect a difference of 25% or more).

Exactly the same pattern emerged for their judgments about who would think "if only" most often. From the short-term perspective, people judged that the actor Joe (62%) would think "if only" more

Table 3
Percentages of Judgments of Who Feels Worse and Who Thinks "If Only" Most Often for Actions and Inactions in the Short- and Long-Terms for Scenarios With Bad Initial States

Question	Short term	Long term
Who feels worse?		
Action	70	76
Inaction	30	24
Who thinks "if only" most often?		
Action	62	61
Inaction	38	39
n	50	51

often than the nonactor Paul (38%), and the difference was reliable (binomial $z = 1.56$, $p < .059$). Likewise from the long-term perspective, people judged that the actor Joe (61%) would think "if only" more often than the nonactor Paul (39%), and the difference was reliable, although somewhat marginal (binomial $z = 1.40$, $p < .08$). There was no difference in the short-term and long-term conditions in the judgments that the actor would think "if only" most (62% vs. 61%), $\chi^2(1, N = 101) = 0.02$, $p > .90$, as Table 3 also shows (and the comparison had sufficient power at the 80% level to detect a difference of 27% or more).

People completed the counterfactual thought by undoing the individual's action or inaction (depending on whether they focused on the actor or the nonactor) in 62% of cases for the short-term scenario and 16% of mutations were general ones (e.g., "I wish I'd taken a chance"). The remaining 22% of completions were non-counterfactual thoughts. None of the undoings focused on the consequences (e.g., "I wish I hadn't lost the money"). Likewise, from the long-term perspective, people completed the thought by undoing the individual's action or inaction in 63% of cases, and 8% of mutations were general. The remaining 29% were non-counterfactual thoughts. Once again, none of the undoings focused on the consequences.

The results show that people mentally undo actions more than inactions from both a long-term and a short-term perspective, even when the information from the outset is bad in the investment scenario, as it was in the college scenario. Of course, it could be argued that a fall in share prices is not as emotionally involving as unhappiness in college, and so the bad initial state is not equivalent in the two sorts of scenarios. Bad investment performance may not be as predictive of future unhappiness as bad college performance. Nonetheless, the findings suggest that the temporal pattern to the agency effect for the college scenario is not the result of its bad outset. The results provide some support for our contention that the crucial difference between the college and investment scenarios lies elsewhere. In the next experiment, we turned from an examination of the bad outset of the scenarios to an examination of the bad outcomes of the scenarios.

Experiment 4: Investments That Are Equally Bad for Actors and Nonactors

We showed that people mentally undid actions even in the long term not only in an investment scenario but also in a college scenario with a good outcome and an investment scenario with a bad outset. The participants switched to mentally undoing inactions from a long-term perspective in situations where the counterfactual consequences from the mentally undone inaction but not the mentally undone action were unknown and possibly better than the real consequences. In all other situations, they mentally undid actions. Our suggestion is that people mentally undo actions because actions are represented explicitly. An alternative explanation for the focus on actions highlights the difference in the subjective outcomes for the actor and the nonactor. In our final experiment, we tested this alternative explanation.

Experiments on the agency effect have relied on scenarios in which the same outcome occurs for the actor and the nonactor (e.g., Gilovich & Medvec, 1994; Gleicher et al., 1990; Kahneman & Tversky, 1982; Landman, 1987). For example, in the investment scenario, a bad outcome (a monetary loss) occurs for both the actor and the nonactor. The outcome is objectively the same for both protagonists, but subjectively it may be subtly different for each one. The bad outcome for the actor results in the presence of a bad consequence, whereas the bad outcome for the nonactor results in the absence of a good consequence; in other words, something bad happens to the actor, whereas something good fails to happen to the nonactor. In the investment scenario, the actor loses money, whereas the nonactor fails to gain money. This subtle difference may lead participants to infer that the actor is in a worse state factually than he was at the outset, whereas the nonactor is in the same neutral state as he was at the outset. This inference may underlie participants' focus on actions and their judgments that actions lead to greater emotional impact and greater frequency of counterfactual thoughts.

We attempted to counter any such inference by ensuring that the outcome was described explicitly as exactly the same for both the actor and the nonactor. In the original investment scenario used in the first experiment, we gave participants the following sort of information for both the actor and the nonactor:

> Paul/Joe saved hard and had £2,000 to invest. . . . At the end of the first year he found out that he would have been better off by £1,000 if he had taken the offer and switched to shares in Company B/if he had turned down the offer and kept his shares in Company B.

In a second scenario, we replaced this sentence with one intended to make the exact nature of the loss more explicit for both the actor and the nonactor:

> Paul/Joe saved hard and had £2,000 to invest. . . . At the end of the first year he found out that his shares had fallen to a value of just £1,000 in Company A and they would have risen to a value of £3,000 if he had taken the offer and switched to shares in Company B/if he had turned down the offer and kept his shares in Company B.

In this explicit-loss version, it is clear for both the actor and the nonactor that the factual situation for each one is that they have lost £1,000 (from a £2,000 investment, their shares have fallen to £1,000), and the counterfactual alternative for each one is that they could have gained £1,000 (from a £2,000 investment their shares would have risen to £3,000). The focus on actions should be eliminated in the new explicit-loss scenario if participants mentally undo actions because they infer that the subjective outcome from the action has led the actor to be in a worse state than when he started—that is, something bad has happened to the actor—

whereas the inaction has led the nonactor to be in the same state as when he started—that is, something good fails to happen to the nonactor.

Method

Materials and design. We constructed a new version of the original investment scenario used in the first experiment, which specified the nature of the outcome more explicitly for both individuals. The new explicit-loss scenario was as follows:

> Paul saved hard and had £2,000 to invest. He bought shares in Company A. During his first year as a shareholder in Company A there was a once-off offer to switch to shares in Company B, but he decided against it. At the end of the first year he found out that his shares had fallen to a value of just £1,000 in Company A and they would have risen to a value of £3,000 if he had taken the offer and switched to shares in Company B.
>
> Joe also saved hard and had £2,000 to invest. He bought shares in Company B. During his first year as a shareholder in Company B there was a once-off offer to switch to shares in Company A, and he decided to do so. At the end of the first year he found out that his shares had fallen to a value of just £1,000 in Company A and they would have risen to a value of £3,000 if he had turned down the offer and kept his shares in Company B.

The original scenario was the same as the one in the first experiment (except that participants were told the protagonists would have been better off by £3,000, not £5,000). The participants again were asked three questions about who would think "if only" most often, how the thought would be completed, and who would feel worse. We gave the original scenario to one group of participants and the explicit-loss version to a second group.

Participants and procedure. The 113 participants were undergraduate students from different departments in the University of Dublin, Trinity College. The 39 women and 74 men, aged 17 to 48 years (M age = 22 years), participated in the experiment voluntarily, and they were assigned at random to one of the two groups. Nine participants were eliminated prior to any data analysis because they failed to complete all three questions. The remaining 104 participants completed all three questions for the original scenario or the explicit-loss scenario (n = 52 in each). They were tested in several groups. They were presented with a booklet that contained the instructions, the scenario and questions, and a debriefing paragraph, each on separate pages. The participants worked at their own pace and provided their answers in the spaces on the sheet.

Results and Discussion

The results showed that participants mentally undid actions in both the original scenario and the explicit-loss scenario, in terms of judgments of greater emotional impact and greater frequency of counterfactual thoughts. For the original scenario, people judged that the actor Joe (75%) would feel worse than the nonactor Paul (25%), and the difference was reliable (binomial $z = 3.47$, $p < .003$). Likewise, for the explicit-loss scenario, people judged that the actor Joe (94%) would feel worse than the nonactor Paul (6%), and the difference was again reliable (binomial $z = 6.24$, $p < .00005$). In fact, people mentally undid actions rather than inactions even more in the explicit-loss scenario than the original (94% vs. 75%), $\chi^2(1, N = 104) = 7.39$, $p > .006$, as Table 4 shows.

A somewhat similar pattern emerged for participants' judgments about who would think "if only" most often. For the original scenario, people judged that the actor Joe (67%) would think "if only" more often than the nonactor Paul (33%), and the difference was reliable (binomial $z = 2.38$, $p < .009$). For the explicit-loss scenario, people judged that the actor Joe (54%) would think "if only" somewhat more often than the nonactor Paul (46%), but the 8% difference between them was not reliable (binomial $z = 0.42$, $p < .39$). Nonetheless, the difference was in the predicted direction, and, most important, there was no difference in the original and explicit-loss scenarios in the judgments that the actor would think "if only" most (67% vs. 54%) $\chi^2(1, N = 104) = 1.97$, $p > .16$, as Table 4 also shows (and the comparison had sufficient power at the 80% level to detect a difference of 27%).

Table 4
Percentages of Judgments of Who Feels Worse and Who Thinks "If Only" Most Often for Actions and Inactions in Scenarios With Explicit-Loss Outcomes

Question	Original	Explicit loss
Who feels worse?		
Action	75	94
Inaction	25	6
Who thinks "if only" most often?		
Action	67	54
Inaction	33	46
n	52	52

People completed the counterfactual thought by undoing the individual's action or inaction (depending on whether they focused on the actor or the nonactor) in 52% of cases for the original scenario. Eight percent of mutations were general (e.g., "I wish I'd taken a chance"), and 2% focused on the consequences. The remaining 38% of completions were noncounterfactual thoughts. Likewise, for the explicit-loss scenario, people completed the thought by undoing the individual's action or inaction in 57% of cases. Six percent of mutations were general and 8% focused on the consequences. The remaining 29% were noncounterfactual thoughts.

The results showed that people mentally undid actions more than inactions for both the original and the explicit-loss scenario. The results suggest that the pervasive focus on actions in the investment scenario did not arise because people inferred that the actor was in a worse state than he was at the outset, whereas the nonactor was in the same state as at the outset. Even when it was made clear that the actor and the nonactor were in an equally worse state than at the outset, participants mentally undid the action more than the inaction. The experiment rules out the alternative explanation of the agency effect.

General Discussion

The results of our experiments show that the temporal pattern to the agency effect occurs only under very specialized circumstances. People mentally undo inactions from the long-term perspective only when the imagined consequences of the mentally undone inaction are unknown and possibly better than the real consequences of the inaction (whereas the imagined consequences for the mentally undone action are known and the same as the real consequences). In all other situations that we have examined, people mentally undid actions. They mentally undid actions even from the long-term perspective when the imagined consequences

for the mentally undone inaction and action were both known and the same (as in the investment scenario in the first experiment) or when the imagined consequences were unknown but not better than the real consequences (as in the college scenario with good outcomes in the second experiment). This pervasive focus on actions in the investment scenario does not seem to arise simply because the initial state for the actor and nonactor is neutral rather than bad, as in the college scenario, and the third experiment ruled out this alternative explanation. This focus also does not seem to arise simply because the subjective outcome seems worse for the actor than the nonactor, and the fourth experiment ruled out this alternative explanation. People mentally undo actions for both bad and good outcomes (Kahneman & Tversky, 1982; Landman, 1987), for scenarios that contain explicit counterfactuals and factual descriptions (Gleicher et al., 1990), and in a variety of real-life situations (e.g., Davis et al., 1995; Gilovich & Medvec, 1994; Gilovich, Medvec, & Chen, 1995).

Our explanation of the agency effect hinges on two simple principles: First, actions are more readily mutated because they are represented explicitly. We propose that actions are represented more explicitly than inactions, by keeping in mind more situations in a mental representation of the events. As a result of the greater amount of information kept in mind about them, actions are more readily available for mutation in the generation of a counterfactual alternative. From a short-term perspective, people mentally undo actions because their mutations are guided by what is represented explicitly (Byrne, 1997).

Second, inactions are more readily mutated from the long-term perspective only when their counterfactual consequences are possibly better than their real consequences (and the action's counterfactual consequences are not). We propose that people can revise their mental representations of events to make additional information explicit. The task of judging who will feel worse and who will think "if only" most about their situation from a long-term perspective requires a consideration of the real and imagined consequences of the events. Actions are more available for mutation if the real and imagined consequences are known for both the actor and the nonactor (as they are in the investment scenario). Actions are more mutable even if the imagined consequences are unknown for the nonactor, once none of the imagined possibilities are better than the real consequences (as in the good-outcome college scenario). However, a reversal to a focus on inactions occurs in a situation, such as the bad-outcome college scenario, in which the real and imagined consequences for the actor are known but only the real consequences are known for the nonactor, and the unknown imagined consequences may possibly be good, that is, better than the real ones for the nonactor. The disparity between the action and inaction in this regard only comes to light from a long-term perspective from which people consider the consequences of events; from a short-term perspective the disparity remains hidden because people focus just on what is explicitly represented in their initial representation of the events. Our account seems to capture most, but of course not all, of the participants responses: A minority of participants, typically up to one third, perform differently from the majority in these tasks. We turn now to a putative sketch of the cognitive processes underlying such a proposal.

Mental Models and the Agency Effect

We suggest that people mentally represent assertions about actions and inactions in different ways by constructing mental models—that is, mental representations of the way the world would be if the assertions were true (Johnson-Laird, 1983). Actions are mentally represented as two separate alternatives: the preaction state of the actor and their postaction state as shown in the following example:

George	Current	Company A
	Past	Company B

where the diagram uses Company A to represent "George has stock in Company A." Separate models are represented on separate lines, and the models are annotated to keep track of their chronological temporal status (see Johnson-Laird, Byrne, & Schaeken, 1992; Schaeken, Johnson-Laird, & d'Ydewalle, 1996). Inactions can be represented more economically by a single mental model: The past state and current state are the same, as follows (Byrne, 1997; Byrne & McEleney, 1997):

Paul	Current and past	Company A

Evidence for such models comes from studies of conditionals, such as the following:

> If George kept his stock in Company B, then it earned $1,200.

A fully explicit set of models represents the three separate situations that capture the way the world would be if the conditional were true (Johnson-Laird & Byrne, 1991):

Company B	$1,200
Not Company B	not $1,200
Not Company B	$1,200

The diagram uses $1,200 to represent "It earned $1,200," and "not" is a propositional-like tag to represent negation (see Johnson-Laird, Byrne, & Schaeken, 1992). However, reasoners construct an initial representation that is more economical than the fleshed-out set, because of the limitations of working memory:

Company B	$1,200
...	

where the three dots represent a model with no explicit content, which captures the idea that alternatives exist that have not been mentally articulated. The models may be "fleshed-out" to be more explicit if necessary (for other details, including a description of the computer program to simulate the theory, see Johnson-Laird & Byrne, 1991, and for recent empirical corroborations, see Byrne, Espino, & Santamaria, 1999; Byrne & Handley, 1997). For example, consider a counterfactual conditional as follows:

> If George had kept his stock in Company B, then he would have been better off by $1,200.

This requires reasoners to represent the information hypothesized, and also the presupposed factual situation:

Factual	Not Company B	Not $1,200
Counterfactual	Company B	$1,200
	...	

As a result of keeping in mind the more explicit representation of the two situations, there are systematic differences in the inferences people make from factual and counterfactual conditionals (Byrne & Tasso, 1999).

The construction of a counterfactual scenario may be constrained by what is explicitly represented in the initial mental representation (e.g., Byrne, 1997; Legrenzi, Girotto, & Johnson-Laird, 1993). Accordingly, it may be easier to construct a counterfactual scenario on the basis of the mental representation of an action; for an inaction, the single economical representation of the current state provides less material from which to construct the alternative possibilities. The longer perspective requires a consideration of the consequences of events (Kahneman, 1995). As a result, people must flesh out their models to be more explicit and to represent explicitly the real and imagined consequences of the actions and inactions. For the investment scenario, the fleshed out models may include the following sort of information:

George	Factual	Company A	not $1,200
	Counterfactual	Company B	$1,200
		...	
Paul	Factual	Company A	not $1,200
	Counterfactual	Company B	$1,200
		...	

An inspection of the set of models shows clearly that each protagonist would have been better off by exactly the same amount had they done things differently. The college scenario is represented in a similar set of models from the short term perspective:

Paul	Factual	College B	Even more unhappy
	Counterfactual	College A	Moderately satisfied
		...	
John	Factual	College A	Even more unhappy
		...	

However, an inspection of the fleshed-out set of models for the college scenario reveals the inequity in information about the counterfactual consequences for the actor and nonactor:

Paul	Factual	College B	Even more unhappy
	Counterfactual	College A	Moderately satisfied
		...	
John	Factual	College A	Even more unhappy
	Counterfactual	College B	**Unknown**
		...	

Reasoners may generate alternative counterfactual possibilities for John:

John	Factual	College A	Even more unhappy
	Counterfactual	College B	**Unhappy**
	Counterfactual	College B	**Happy**
		...	

When the outcome of the scenario is good, the counterfactual consequences are unknown for the nonactor but might be worse than the real consequences:

Paul	Factual	College B	Happy
	Counterfactual	College A	Moderately satisfied
		...	
John	Factual	College A	Happy
	Counterfactual	College B	**Unhappy**
	Counterfactual	College B	**Happy**
		...	

This account is a preliminary sketch of a theory of counterfactual mutability. It is necessarily tentative, is based on already existing data, and requires empirical test.

Alternative Explanations

We have attempted to sketch some of the cognitive factors that may contribute to the mutability of actions. Motivational variables may also be important, but the results of our experiments suggest that motivational variables may not always play a central role. Consider, for example, the seven mechanisms outlined by Gilovich and Medvec (1995a), five motivational and two cognitive, that may contribute to the temporal pattern of the focus on actions and the reversal to a focus on inactions. The five motivational variables should continue to operate to create a focus on inactions in the long term for the investment scenario examined in the first experiment. The three motivational mechanisms that reduce regret for actions—taking compensatory steps, looking for silver linings, and engaging in more dissonance reduction—should apply equally to the actor in the investment scenario as in the college scenario. Two motivational mechanisms to increase regret for inactions also apply equally to the investment scenario, the inaction may be inexplicable because of an increase in retrospective confidence in the action's success and the nonsalience of variables that inhibited the action. These five motivational mechanisms predict a focus on inactions in the long-term investment scenario and so they are inconsistent with the results, for example, of the first experiment. The action effect in the long term was just as strong as the action effect in the short term for the investment scenario in the first experiment, which suggests that people judge that the actor and the nonactor will not make their peace with their financial loss over time (Kahneman, 1995). Instead cognitive mechanisms, in particular the unknown nature of the counterfactual consequences of some inactions, bounded only by the limits of the imagination of what might have been, may be critical (Gilovich & Medvec, 1995b; Kahneman, 1995).

For the good-outcome scenarios examined in the second experiment, it is also not clear that a motivational explanation would provide a viable account. For example, of the seven mechanisms that Gilovich and Medvec (1995a) outlined, the first three, which reduce the emotional impact of regrettable actions, do not apply to feeling better about actions: There is no need to engage in compensation, silver linings, or dissonance. Hence, no reversal of an action effect to an inaction effect over time would be expected for good outcomes. But, two motivational factors do apply to the good-outcome scenarios: Emotions for inactions may be enhanced by their inexplicability because of retrospective confidence, and inhibitory factors may be less accessible. Moreover, the cognitive factors also apply to the good-outcome scenarios: Inactions may be more available because they are open-ended (their consequences are unknown) and thus memorable, and their counterfactual consequences are potentially infinite. Hence, at the very least, this account would predict a reduction, even a reversal of the focus on actions over time for good outcomes. The results of the second experiment show that such a reversal does not occur and so they are difficult for the motivational account to explain. The results of the four experiments reported here suggest that cognitive variables are centrally important in an explanation of the mutability of actions.

Conclusions

Our experiments showed that the temporal pattern to the agency effect occurs only in a limited set of circumstances. The experiments showed no temporal pattern to the agency effect when the counterfactual consequences from the mentally undone inaction were known, no temporal pattern even when the counterfactual consequences from the mentally undone inaction were unknown once the outcome of the situation was good, no temporal pattern even when the situation for both the actor and the nonactor was specified to be bad from the outset, and a continued focus on actions even when the nature of the loss was specified to be equal for both the actor and the nonactor.

In real life, the counterfactual consequences from people's inactions are sometimes knowable, for example, in situations where the outcome depends largely on external factors such as the success of the stock exchange; sometimes the consequences are unknowable, for example, in situations where the outcome depends largely on internal factors such as the success of an individual's adjustment to college. Our experiments suggest that it is only in situations where the counterfactual consequences of mentally undone inactions are unknown—and possibly better than the factual consequences—that there is a temporal pattern to the agency effect. Our results show that where the counterfactual consequences of mentally undone actions and inactions are both known and similar, there is no temporal pattern to the agency effect: People judge that actions elicit most emotion and most counterfactual thought generation from both a short-term and a long-term perspective.

We have suggested that the action effect can be explained with reference to the nature of the mental representations that people construct of factual situations. People represent actions more explicitly than inactions, and they can flesh out their representation to be more explicit if need be to consider the counterfactual consequences of mentally undone actions and inactions. Our account relates the cognitive processes underlying the construction of counterfactual scenarios to the cognitive processes that underlie other aspects of high-level cognition, including deductive reasoning (Johnson-Laird & Byrne, 1991) and naive judgments of probability (Johnson-Laird, Legrenzi, Girotto, Sonino Legrenzi, & Caverni, 1999). It suggests that the mutability of actions and inactions depends crucially on differences in the way in which they are mentally represented.

References

Byrne, R. M. J. (1997). Cognitive processes in counterfactual thinking about what might have been. In D. L. Medin (Ed.), *The psychology of learning and motivation, Vol. 37* (pp. 105–154). San Diego, CA: Academic Press.

Byrne, R. M. J., Espino, O., & Santamaria, C. (1999). Counterexamples and the suppression of inferences. *Journal of Memory and Language, 40,* 347–373.

Byrne, R. M. J., & Gillanders, R. (1999). Explicit counterfactuals and the agency effect for good outcomes. *Unpublished manuscript.*

Byrne, R. M. J., & Handley, S. J. (1997). Reasoning strategies for suppositional deductions. *Cognition, 62,* 1–49.

Byrne, R. M. J., & McEleney, A. (1997). Cognitive processes in regret for actions and inactions. In M. Shafto & P. Langley (Eds.), *Proceedings of the 19th Annual Conference of the Cognitive Science Society,* (pp. 73–78). Hillsdale, NJ: Erlbaum.

Byrne, R. M. J., Segura, S., Culhane, R., Tasso, A., & Berrocal, P. (2000). The temporality effect in counterfactual thinking about what might have been. *Memory & Cognition, 28,* 264–281.

Byrne, R. M. J., & Tasso, A. (1999). Deductions from factual, possible and counterfactual conditionals. *Memory & Cognition, 27,* 726–740.

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.

Dunning, D., & Parpal, M. (1989). Mental addition versus subtraction in counterfactual reasoning: On assessing the impact of personal actions and life events. *Journal of Personality and Social Psychology, 57,* 5–15.

Garnham, A. (1985). *Psycholinguistics: Central topics.* London: Methuen.

Gilovich, T., & Medvec, V. H. (1994). The temporal pattern to the experience of regret. *Journal of Personality and Social Psychology, 67,* 357–365.

Gilovich, T., & Medvec, V. H. (1995b). Some counterfactual determinants of satisfaction and regret. In N. J. Roese & J. M. Olson (Eds.), *What might have been: The social psychology of counterfactual thinking* (pp. 259–282). Hillsdale, NJ: Erlbaum.

Gilovich, T., & Medvec, V. H. (1995a). The experience of regret: What, when, and why. *Psychological Review, 102,* 379–395.

Gilovich, T., Medvec, V. H., & Chen, S. (1995). Commission, omission, and dissonance reduction: Coping with regret in the 'Monty Hall' problem. *Personality and Social Psychology Bulletin, 21,* 182–190.

Gilovich, T., Medvec, V. H., & Kahneman, D. (1998). Varieties of regret: A debate and partial resolution. *Unpublished manuscript.*

Ginsberg, M. L. (1986). Counterfactuals. *Artificial Intelligence, 30,* 35–79.

Girotto, V., Legrenzi, P., & Rizzo, A. (1991). Event controllability in counterfactual thinking. *Acta Psychologica, 78,* 111–133.

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,* 284–295.

Hofstadter, D. R. (1985). *Metamagical themas: Questing for the essence of mind and pattern.* London: Penguin.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, England: Cambridge University Press.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction.* Hillsdale, NJ: Erlbaum.

Johnson-Laird, P. N., Byrne, R. M. J., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review, 99,* 418–439.

Johnson-Laird, P. N., Legrenzi, P., Girotto, V., Sonini Legrenzi, M., Caverni, J. -P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review, 106,* 62–88.

Kahneman, D. (1995). Varieties of counterfactual thinking. In N. J. Roese & J. M. Olson (Eds.), *What might have been: The social psychology of counterfactual thinking* (pp. 375–396). Hillsdale, NJ: Erlbaum.

Kahneman, D., & Miller, D. (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.

Keane, M. T. (1997). What makes an analogy difficult? The effects of order and causal structure in analogical mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 946–967.

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.

Legrenzi, P., Girotto, V., & Johnson-Laird, P. N. (1993). Focusing in reasoning and decision-making. *Cognition, 49,* 37–66.

McCloy, R., & Byrne, R. M. J. (in press). Counterfactual thinking about controllable actions. *Memory & Cognition.*

Miller, D. T., & Gunasegaram, S. (1990). Temporal order and the per-

ceived mutability of events: Implications for blame assignment. *Journal of Personality and Social Psychology, 59,* 1111–1118.

Miller, G., & Johnson-Laird, P. N. (1976). *Language and perception.* Cambridge, England: Cambridge University Press.

N'Gbala, A., & Branscombe, N. R. (1997). When does action elicit more regret than inaction and is counterfactual mutation the mediator of this effect. *Journal of Experimental Social Psychology, 33,* 324–343.

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

Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin, 121,* 133–148.

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

Savitsky, K., Medvec, V. H., & Gilovich, T. (1997). Remembering and regretting: The Zeigarnik effect and the cognitive availability of regrets for action and inaction. *Personality and Social Psychology Bulletin, 23,* 248–257.

Schaeken, W., Johnson-Laird, P. N., & d'Ydewalle, G. (1996). Mental models and temporal reasoning. *Cognition, 60,* 205–234.

Shafir, E., & Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive Psychology, 24,* 449–474.

Spellman, B. A. (1997). Crediting causality. *Journal of Experimental Psychology: General, 126,* 323–348.

Wells, G. L., & Gavanski, I. (1989). Mental simulation of causality. *Journal of Personality and Social Psychology, 55,* 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.

Zeelenberg, M. (1996). *On the importance of what might have been: Psychological perspectives on regret and decision making.* Unpublished doctoral dissertation, University of Amsterdam.

Received August 28, 1998
Revision received February 10, 2000
Accepted March 1, 2000 ■