Journal of Personality and Social Psychology
1995, Vol. 69, No. 6, 1069–1086

0022-3514/95/$3.00

Multiple Explanation: A Consider-an-Alternative Strategy for Debiasing Judgments

Edward R. Hirt and Keith D. Markman
Indiana University Bloomington

Previous research has suggested that an effective strategy for debiasing judgments is to have participants "consider the opposite." The present research proposes that considering any plausible alternative outcome for an event, not just the opposite outcome, leads participants to simulate multiple alternatives, resulting in debiased judgments. Three experiments tested this hypothesis using an explanation task paradigm. Participants in all studies were asked to explain either 1 hypothetical outcome (single explanation conditions) or 2 hypothetical outcomes (multiple explanation conditions) to an event; after the explanation task, participants made likelihood judgments. The results of Studies 1 and 2 indicated that debiasing occurred in all multiple explanation conditions, including those that did not involve the opposite outcome. Furthermore, the findings indicated that debiased judgments resulted from participants' spontaneous consideration of additional alternatives in making their likelihood judgments. The results of Study 3 also identified the perceived plausibility of the explained alternative as an important moderating variable in debiasing.

We all spend a great deal of time thinking about the future. How often have you found yourself considering questions like the following: Will Bob Dole be the presidential candidate for the Republican party in 1996? Will the San Francisco 49ers repeat as Super Bowl champs in 1996? Will investment in this real estate venture pay off and allow me to retire early? Will the decision to start a family bring me and my spouse closer together?

In answering questions such as these, one must project into the future and predict whether the specified outcome is likely to occur. However, there is a considerable amount of research (cf. Johnson & Sherman, 1990; Koehler, 1991) demonstrating that merely specifying a particular future event or outcome to think about leads people to subsequently perceive that event or outcome as more likely. One of the clearest demonstrations of this effect is the *explanation bias.* In this work, participants have been asked to imagine or generate explanations for hypothetical future events and outcomes. Results consistently demonstrate that participants asked to imagine or explain how a hypothetical outcome might be true show increased subjective likelihood estimates for the target outcome relative to participants not given the imagination–explanation task. For example, in the initial demonstration of the explanation bias, Ross, Lepper, Strack, and Steinmetz (1977) had participants read detailed clinical case histories of psychiatric patients, and participants were asked to write explanations for why particular events (e.g., committing suicide or contributing to the Peace Corps) might have occurred later in the patient's life. Even though the event to be explained was known to be hypothetical, participants who had explained a given event believed that the patient was actually more likely to perform these behaviors in the future.

Subsequent experiments have demonstrated the explanation bias to be a robust phenomenon. The explanation bias has been shown in a number of different domains, ranging from political elections (Carroll, 1978) to sporting events (Hirt & Sherman, 1985; S. J. Sherman, Zehner, Johnson, & Hirt, 1983) to social theories (Anderson, Lepper, & Ross, 1980). In addition, research indicates that explanation biases can occur even when one is explaining hypothetical future events involving oneself (Campbell & Fairey, 1985; S. J. Sherman, Skov, Hervitz, & Stock, 1981). S. J. Sherman et al. (1981) found that participants who explained hypothetical success on a future anagram task believed that success was more likely in the future. Moreover, these changes in likelihood estimates after explanation tasks appear to be quite persistent and resistant to change. Anderson (1983b) found that participants continue to exhibit increased confidence in the validity of explanation-induced social theories a week after the explanation task. Furthermore, explanation biases have been shown to affect not only likelihood estimates but subsequent evaluations (Anderson & Sechler, 1986) and future behavior as well (Gregory, Cialdini, & Carpenter, 1982; R. T. Sherman & Anderson, 1987; S. J. Sherman et al., 1981).

Explanations for the Explanation Bias

Given the robustness and ubiquity of the explanation bias, it is important to understand why this bias occurs. The dominant explanation offered for the explanation bias evokes the opera-

Edward R. Hirt and Keith D. Markman, Department of Psychology, Indiana University Bloomington. Keith D. Markman is now at the Department of Psychology, Ohio State University.

This research was supported in part by a biomedical research support grant from the University of Wisconsin. We wish to thank Russ Fazio, Frank Kardes, and Jim Sherman for their helpful comments on a draft of this article. We also wish to thank Ivan Amodt, George Bizer, Allan Bogan, Brian Cawley, Martina Gross, Heather Huffman, Craig Meyer, Jamie Pierce, and Pat Wells for their assistance in serving as experimenters and coders for the reported studies.

Correspondence concerning this article should be addressed to Edward R. Hirt, Department of Psychology, Indiana University, Bloomington, Indiana 47405.

tion of the availability heuristic (Tversky & Kahneman, 1973). According to this judgmental principle, people judge the likelihood of future events or outcomes on the basis of the ease with which instances and examples can be brought to mind. Events whose instances readily come to mind are judged as more likely than events whose instances do not readily come to mind (cf. Tversky & Kahneman, 1973). Like the other judgmental heuristics, the availability heuristic is a reasonable rule of thumb that leads to relatively accurate judgments most of the time; however, errors in judgment will occur to the extent that participants fail to recognize that other factors (independent of event frequency) have affected the ease of retrieval of instances. Because an explanation task makes information consistent with an outcome more salient and accessible, causal arguments consistent with that outcome will be more readily and easily retrieved at the time of judgment than arguments consistent with alternative outcomes (cf. Anderson et al., 1980; Anderson, New, & Speer, 1985). As a result, judgments will be biased in the direction of the outcome explained.

Recently, Koehler (1991) presented a slightly different explanation for the explanation bias. According to Koehler, an explanation task draws attention to a single, specified hypothesis (which he labeled a *focal hypothesis*). The establishment of a focal hypothesis then prompts the person to adopt a *conditional reference frame* in which the focal hypothesis is temporarily assumed to be true (cf. Gilbert, 1991). The person then evaluates all of the relevant evidence within this reference frame and judges how well the hypothesis fits the available evidence. However, adopting a conditional reference frame leads to systematic biases in information search and the interpretation of evidence in the direction of the focal hypothesis. Thus, people are more likely to perceive a good fit between the evidence and the focal hypothesis, resulting in overestimation of the likelihood of the focal hypothesis.

Both of these explanations argue that an explanation task enhances the accessibility of information consistent with the event explained at the time of judgment. Because participants base their judgments on the information accessible at the time of judgment and do not take into account the biasing effects of the earlier explanation task, their likelihood judgments are systematically biased in favor of the event explained. Participants fail to consider how well the evidence might fit alternative outcomes, much as people engage in truncated searches for the causes of past events (Shaklee & Fischhoff, 1982). Indeed, these same processes have been used to explain many other judgmental biases, including overconfidence (Lichtenstein, Fischhoff, & Phillips, 1982; Vallone, Griffin, Lin, & Ross, 1990), hindsight biases (Fischhoff, 1975; Fischhoff & Beyth, 1975), and biases in hypothesis testing (Devine, Hirt, & Gehrke, 1990; Skov & Sherman, 1986; Snyder & Swann, 1978).

Debiasing the Effects of Explanation Tasks

Given that explanation tasks result in biased and suboptimal judgments, several researchers have explored debiasing techniques for explanation effects. The primary debiasing technique that takes this approach has been that of having participants engage in a counterexplanation task. Participants asked to explain one outcome of an event are then asked to explain an alternative outcome to the same event. Because most of the tasks used in experiments have involved dichotomous outcomes, most counterexplanation tasks have taken the form of a "consider-the-opposite" condition. Several experiments have demonstrated that consider-the-opposite instructions reduce the explanation bias (Anderson, 1982; Anderson & Sechler, 1986; Hoch, 1984; Lord, Lepper, & Preston, 1984). Indeed, this same technique has proven successful in reducing other judgmental errors such as overconfidence (Hoch, 1985; Koriat, Lichtenstein, & Fischhoff, 1980), hindsight bias (Fischhoff, 1982; Slovic & Fischhoff, 1977), and confirmatory hypothesis testing (Wason, 1960; Wason & Golding, 1974).[1]

Why is a counterexplanation task effective in debiasing judgments? The preferred explanation for the debiasing effects of counterexplanation tasks is that such tasks explicitly force participants to consider alternative outcomes for an event. That is, participants must generate evidence that supports an alternative outcome (and thus contradicts the previously explained outcome), resulting in a more balanced and objective evaluation of the relevant evidence at the time of judgment. Koehler (1991) has argued that the counterexplanation task "breaks the inertia" that sets in once a given frame is adopted. In addition, Arkes (1991) argued that counterexplanation strategies prime stimuli other than the ones that would normally be accessed. Thus, presumably, considering the opposite leads people to use a more thorough, qualitatively better judgmental process in making their likelihood estimates.

Although this explanation for the effects of counterexplanation tasks is both intuitively plausible and appealing, it focuses largely on metaphoric descriptions rather than specifying the processes that underlie the debiasing of judgments. Indeed, an examination of the extant studies using counterexplanation tasks reveals that little attention has been given to collecting process data aimed at isolating the source of debiasing; in most studies, only judgment data have been assessed. As a result, these studies are open to a number of different explanations. In this article, we outline several possible explanations for the debiasing effects of counterexplanation tasks and describe a series of studies we conducted to critically test and evaluate each of these explanations.

Potential Processes Underlying the Debiasing Effects of Counterexplanation Tasks

In this section, we consider three possible processes by which counterexplanation tasks might produce unbiased judgments. First, participants might illustrate unbiased judgments based on the use of a simple averaging rule. That is, given that the counterexplanation task has made arguments favoring the opposite outcome more accessible in memory (cf. Anderson & Godfrey, 1987; Anderson et al., 1985), the availability heuristic predicts that counterexplanation participants should be able to bring to mind easily arguments favoring both outcomes; thus, they should believe that either one of these outcomes could oc-

[1] Another conceptually similar variant of this counterexplanation technique, an "inoculation condition" in which participants are asked to give an explanation for both possible outcomes before being presented with a set of information, has also proven to be effective at debiasing participants' judgments (Anderson, 1982).

cur. However, on the judgment task, they are forced to make a choice regarding the likely future outcome. As a result, they may simply average or "split the difference" between the implications of the two opposing explanation tasks in arriving at their likelihood judgments.[2] Indeed, a split-the-difference strategy would appear to be a reasonable heuristic for participants to use when faced with such a judgment task; thus, it constitutes a viable explanation for the debiasing effects of counterexplanation tasks.

A second explanation argues that debiasing is the result of increasing participants' uncertainty. Given that the participant is asked to explain two opposing outcomes to the same event, he or she might be completely confused as to what to predict on the subsequent judgment task. When the participant is then forced to make a choice on a bipolar judgment scale (anchored at *event X will definitely occur* and *event Y will definitely occur*), he or she may feel compelled to simply endorse the middle (or indifference point) of the scale, reflecting uncertainty. Thus, instead of reflecting the enhanced availability of arguments supporting both sides and true ambivalence about the outcome of the event (arguments consistent with the split-the-difference explanation), this explanation argues that debiasing reflects a default response indicating uncertainty. Indeed, with any judgmental scale, it is unclear whether a response at the midpoint of the scale reflects a genuine response indicating ambivalence or a default response indicating uncertainty (see Kaplan, 1972, and Thompson, Zanna, & Griffin, in press, for a related argument with regard to responses to attitudinal scales).

Alternatively, we propose a third process. A counterexplanation task presents participants with a task of considering one alternative outcome for the event in question: Participants must undo their prior explanation for the event and construct an explanation supporting a different outcome. We hypothesize that successful completion of the counterexplanation task should lead participants to realize that the outcome of the event is not as predictable as previously believed. This realization may then lead participants to consider additional alternatives (beyond those specified in the explanation and counterexplanation tasks) in making their likelihood judgments. Thus, we argue that, in answering the judgment question, counterexplanation participants do not simply consider only the two explained alternatives (as the split-the-difference view would predict). Instead, participants are presumed to use the *simulation* heuristic and to engage in multiple simulation runs of the potential outcomes of the event. According to this judgmental heuristic, participants judge the likelihood of an outcome on the basis of the ease (and frequency) with which scenarios leading to a particular outcome can be constructed (Kahneman & Tversky, 1982); outcomes that are easily simulated are judged as relatively likely, and outcomes more difficult to simulate are judged as relatively unlikely (cf. S. J. Sherman, Cialdini, Schwartzman, & Reynolds, 1985). This same process has been argued to underlie counterfactual reasoning of past events (Kahneman & Miller, 1986; Wells & Gavanski, 1989); thus, we are proposing that the same process occurs in the prediction of future outcomes following a counterexplanation task. Specifically, we argue that to the extent that scenarios consistent with the alternative outcome are also found to be easy to simulate, counterexplanation participants consider additional alternatives (to those specified in the explanation and counterexplanation tasks) and engage in additional simulation runs for these alternatives. In this way, the counterexplanation task, to use Koehler's (1991) terminology, breaks the inertia in considering alternatives, leading participants to consider a fuller, more complete set of alternatives as they judge the likely outcome of the event. Thus, according to this explanation, debiasing results from the fact that the simulation of additional alternative outcomes for the event reveals that plausible alternatives to the initially explained outcome exist. Participants then base their likelihood judgments on the results of the simulation process, judging the most easily simulated outcome to the event as most probable.

The Present Research

The present research was designed to provide more conclusive evidence regarding the process by which counterexplanation tasks affect likelihood judgments. In particular, the present studies attempted to provide direct evidence in support of the notion that counterexplanation tasks engage the use of the simulation heuristic. In addition, the studies attempted to rule out the two rival hypotheses (split-the-difference and uncertainty) for the effects of counterexplanation on judgment.

Given that all three of these processes serve as viable explanations for the debiasing effects of counterexplanation, how might one distinguish among them? In the service of this goal, the studies reported here included events for which more than two possible outcomes are considered. Indeed, the use of events involving dichotomous outcomes in which participants explain one outcome and its opposite hopelessly confounds the implications of these different processes. However, when one considers events with multiple potential outcomes, the implications of these various strategies for the judgment data diverge. Consider the example of trying to predict the academic success of a special education participant mainstreamed into an elementary school classroom. Participants might explain a positive change, no change, or a negative change in the participant's academic performance. A split-the-difference process predicts that judgments will be an average of the two component (i.e., explained) outcomes; thus, counterexplanation participants explaining first a positive change and then no change should come out with a more positive net judgment than participants explaining a negative change and then no change. An uncertainty process, however, predicts that participants' judgments will correspond to the middle (indifference point) of the judgmental scale, re-

[2] Indeed, there are a number of motivations that might underlie participants' use of a "split-the-difference" strategy. In this article, we focus on an availability heuristic argument behind the use of this strategy. However, one could also argue that participants might split-the-difference as a result of experimental demand. Although the specific motivations behind the use of the split-the-difference strategy are beyond the scope of this article, we refer interested readers to an article by Lord et al. (1984) in which the authors argued extensively and provided data against a demand interpretation of the effects of consider-the-opposite tasks. However, it is worth noting that a demand argument cannot account for the pattern of results obtained across the three studies reported here. Thus, in the present context, we are referring only to an availability heuristic explanation for the split-the-difference strategy.

gardless of which two outcomes are explained; thus, participants explaining a positive change and then no change should render the same judgment (anchored at the midpoint of the scale) as participants explaining a negative change and then no change. A simulation heuristic process predicts that judgments would reflect the outcome for which scenarios consistent with that outcome are most easily constructed. As with the uncertainty argument, this view predicts that participants explaining a positive change and then no change and participants explaining a negative change and then no change should make the same judgments. However, unlike the uncertainty argument, these judgments need not necessarily reside at the middle of the scale (unless the outcome of no change is the most easily simulated outcome); instead, judgments should correspond to the most easily and readily simulated outcome of the event. The differences in the predictions of each of the three processes for the judgment data (for events with multiple outcomes) are represented in Table 1.

In addition to the judgment data, a second means by which one can discriminate among these processes is by examining participants' confidence in their judgments. An uncertainty argument would predict that counterexplanation participants feel confused about what to predict and would exhibit significantly lower judgmental confidence in their estimates (as compared with single explanation participants). The split-the-difference argument would anticipate that participants' judgmental confidence should remain high because counterexplanation participants, like single explanation participants, are simply basing their judgments on the set of arguments currently available in memory. Likewise, the simulation heuristic view predicts that confidence should be high because participants are considering multiple alternatives (and thus are using more information) in making their judgment of the probable outcome of the event.[3] Thus, in Studies 1 and 2 of the present research, participants were asked to express their confidence in their likelihood judgments. Predictions regarding the confidence measure are also represented in Table 1.

A third source of information that could potentially discriminate among these different processes is an examination of participants' justification of their judgments. In one study (Study 1), participants were asked to write down the reasons for making their predictions. These reasons could then be content coded for both (a) expressions of consideration of different possible outcomes in rendering their judgments and (b) expressions of (un)certainty and (lack of) confidence in justifying their judgments. An uncertainty view would predict that counterexplanation participants' justifications would be fraught with expressions of uncertainty. A split-the-difference view would predict that counterexplanation participants would express with confidence consideration of both explained outcomes in their justifications. The simulation heuristic view would also predict that counterexplanation participants would confidently express their judgments and should express consideration of multiple alternative outcomes; however, the predictions for the simulation heuristic view differ from those of the split-the-difference view in that the alternative outcomes are just as likely to be ones different from those previously explained (during the explanation tasks) as ones considered in prior explanations (see Table 1).

Study 1

Study 1 was designed to provide an initial test of the notion that counterexplanation participants are using the simulation heuristic and consider multiple alternatives in making their likelihood estimates. In addition, Study 1 attempted to rule out two rival (i.e., split-the-difference and uncertainty) explanations for the debiasing effects of counterexplanation. Participants were asked to consider the relationship between the personality dimension of riskiness versus conservatism and success as a fire fighter (cf. Anderson, 1982; Anderson et al., 1980).[4] All participants were presented with the responses of two fire fighters, one successful and one unsuccessful, to a test assessing preferences for risky versus conservative actions. Participants in the single explanation conditions were then asked to explain one of three possible relationships: a positive relationship between riskiness and success as a fire fighter, a negative relationship, or no relationship. Participants in the multiple (i.e., counterexplanation) conditions explained one relationship and then were asked to explain a second alternative relationship. All possible combinations of counterexplanation conditions and order of explanation were included, resulting in six different multiple explanation conditions (positive–no, positive–negative, negative–no, negative–positive, no–positive, and no–negative). After the explanation task(s), all participants made judgments about the actual relationship between riskiness and success as a fire fighter, indicated their confidence in their judgments, and explained the reasons underlying their judgments. A no explanation control condition was also included for comparison purposes (see Koehler, 1991, for a discussion of the importance of including such control conditions in explanation experiments).

We expected that participants in the single explanation conditions would show the explanation bias, replicating previous work (Anderson, 1982). Moreover, we expected that participants in the multiple explanation conditions would show evidence of debiasing the effects of the initial explanation task. In addition, consistent with the simulation heuristic view, we expected that all multiple explanation conditions would be equally effective. Thus, contrary to the averaging or split-the-difference explanation, the positive–no and negative–no multiple explanation conditions should debias participants' judgments to the same extent as the positive–negative (i.e., con-

[3] It is important to note that these predictions regarding confidence data refer to confidence in one's *judgment* rather than confidence in the *focal hypothesis.* We argue that participants using the split-the-difference or simulation heuristic strategy would express confidence in the judgments they make concerning the future outcome of the event; however, we acknowledge that the counterexplanation task should undermine participants' confidence in the focal hypothesis because it prompts the consideration of evidence to the contrary. That is, participants using either of these strategies may believe the focal hypothesis to be less likely but should still have confidence in the specific judgment they have made. Given that, in Studies 1 and 2, we measured confidence in terms of participants' confidence in their judgments, we focus on these predictions at present. However, we return to this issue in Study 3, in which we included a measure of participants' confidence in the focal hypothesis.

[4] We would like to thank Craig Anderson for providing us with the materials for this experiment.

Table 1
Predictions of the Three Different Processes Potentially Underlying the Debiasing Effects of Counterexplanation Tasks

Prediction	Split the difference	Uncertainty	Simulation heuristic
Underlying process	Counterexplanation task increases availability of counterarguments to initially explained outcome	Participants become confused and uncertain about what to predict on judgment task	Participants spontaneously consider and simulate additional alternatives
Judgment data	Participants should use averaging rule; judgments midway between the two explained outcomes	Participants provide default response at indifference point of the judgment scale	Judgments should reflect the most easily simulated outcome (paralleling no explanation control participants)
Confidence data	Confidence should be **high**; judgments based on available arguments	Participants should express low confidence	Participants should express **high** confidence
Reasons data	Reasons should reflect consideration of both explained outcomes	Reasons should be expressed with much **hedging** and uncertainty (qualifiers)	Reasons should express consideration of multiple possible outcomes, **including outcomes not previously explained**

sider-the-opposite) condition. In turn, judgments in all multiple explanation conditions should not differ from those in the no explanation control condition, indicating that the biasing effects of the prior explanation task have been completely overcome. Finally, consistent with the simulation heuristic view, we expected to see no differences in judgmental confidence in multiple explanation conditions (relative to single or no explanation conditions).

Method

Participants

Participants were 326 participants enrolled in introductory-level psychology courses at Indiana University. Participants received course extra credit for their participation. Participants were tested in groups ranging in size from 2 to 25. Data from 16 participants were dropped because these participants failed to explain the proper event(s). Three additional participants were dropped because they failed to complete the dependent measures. Thus, data from 307 participants remained for inclusion in the subsequent analyses.

Procedure

In accord with the procedure used by Anderson (1982), participants were told that the experiment concerned how well people are able to discover relationships between personality traits and behavior. In particular, they would be asked to consider the relationship between eventual success or failure as a fire fighter and the general tendency toward riskiness or conservatism. Participants were told that they would be given case history information about one highly successful and one unsuccessful fire fighter. In addition to nondiagnostic background information, participants received detailed information about the success or failure of the two target fire fighters (as assessed by superiors' ratings) and four of each fire fighter's most representative responses to the "risky–conservative choice test," a paper-and-pencil test designed to measure preferences for risky versus conservative actions (and purportedly given when the fire fighters entered the training program). Participants were asked to read over the information carefully and to try to get a sense of the riskiness versus conservatism of each fire fighter's responses. Participants were given 8 min to examine the case history information.

All participants were then presented with the same set of information. In actuality, two of the four responses for each fire fighter were risky and two were conservative. Thus, there was no objective relationship in the presented data between riskiness–conservatism and success–failure as a fire fighter. We chose to present a balanced set of data so that participants could find evidence in support of any relationship (positive, negative, or no relationship) between riskiness and success as a fire fighter in these data.

Explanation tasks. After examination of the case history information, experimental participants were given an explanation task. Control (no explanation) participants were given a filler task (math problems) to do to equate the amount of time between examination of the case history information and completion of the dependent measures. Three different explanation tasks were included in this study. In the positive relationship condition, participants were asked to imagine that the relationship between riskiness and success as a fire fighter was a positive one (i.e., high riskiness is associated with high ability or success as a fire fighter). Participants in the negative relationship condition were asked to imagine that the relationship was a negative one (i.e., high riskiness is associated with low ability or failure as a fire fighter). Participants in the no relationship condition were asked to imagine that there was no relationship between the two variables (i.e., personal riskiness or conservatism is unrelated to ability or success as a fire fighter). In all conditions, participants were asked to write down those factors or reasons that might help one to explain how or why that relationship might be true.

At this point, the procedure followed by participants in the single and multiple explanation conditions differed. Participants in the multiple explanation conditions read the following instructions:

> As psychologists, we are interested in how people explain hypothetical relationships between variables. Your previous task was to imagine and explain a particular relationship between these variables. Obviously, this is not the only possible relationship between these two variables. Now we'd like you to imagine and explain a *different* relationship. On the following page, you will find a new explanation task. Please approach this task as if you were doing it for the first time. Once again, you should write out a complete explanation for this new relationship between these variables on the following page.

These participants then went on to complete a second, different explanation task. As a result, six multiple explanation conditions were created: positive–no, positive–negative, no–positive, no–negative, negative–positive, and negative–no. In place of the second explanation task, participants in the single explanation conditions (positive only, negative only, and no only) completed a filler task (math problems) to equate the amount of time between the examination of the case history information and the completion of the dependent measures. Control participants, as well, completed a second set of math problems in place of the second explanation task.

Dependent measures. After the explanation task(s) and filler task(s), all participants completed the dependent measures. The primary measure consisted of participants' rating of their own personal beliefs about the relationship between riskiness and success in fire fighting. Participants were presented with a 13-point scale ranging from *highly negative relationship* (1) to *highly positive relationship* (13; 7 was labeled *no relationship*). A second measure asked participants to estimate the actual correlation between riskiness and success as a fire fighter (range = −1 to 1). A third measure asked participants to report their confidence in their estimated relationship on a 13-point scale ranging from *not at all confident* (1) to *very confident* (13). Finally, participants were asked to briefly explain the reasons for their estimates and to comment on any personal experiences or relationships that might have affected their responses.

Results

Single Explanation Conditions

Table 2 presents the data obtained for all conditions on the main dependent measures. Our primary interest in this study was to investigate the debiasing effects of different counterexplanation tasks on likelihood estimates. However, before we could examine the debiasing effects in the multiple explanation conditions, we needed to demonstrate that the explanation bias occurred in the single explanation conditions. Indeed, a one-way analysis of variance (ANOVA) on the estimated relationship for the three single explanation conditions revealed a significant main effect of explanation condition, $F(2, 98) = 34.58$, $p < .001$. Participants in the negative only condition estimated a significantly more negative relationship between riskiness and success as a fire fighter (M = 5.87) than did participants in the no only (M = 8.10) and positive only (M = 10.12) conditions, both ts > 3.80, ps < .001. In addition, participants in the positive only condition estimated a significantly more positive relationship than did participants in the no only condition, $t(62) = 5.08$, $p < .001$.

Similar results were obtained on the estimated correlation coefficient measure, $F(2, 96) = 35.54$, $p < .001$. Participants in the negative only condition estimated a mean correlation of −.17 between riskiness and fire fighting ability, as compared with correlations of .29 in the no only condition and .65 in the positive only condition, all ts > 4.00, ps < .001. However, no significant differences in participants' confidence ratings were obtained, $F(2, 98) = 1.33$, *ns*.

Comparisons with the control condition further attested to the effectiveness of the explanation manipulations. Participants in the negative only condition estimated a significantly more negative relationship between riskiness and fire fighting ability than did control participants (M = 8.24), whereas participants in the positive only condition estimated a significantly more positive relationship, both ts > 3.99, ps < .001. Estimates of participants in the control and no only conditions did not differ, $t(80) = 0.26$, *ns*. Comparable results were obtained with the mean correlation coefficient: The mean correlation estimated by control participants (M = .32) differed significantly from estimates of participants in both the negative only and positive only conditions, both ts > 3.95, ps < .001, but did not differ significantly from the estimate of participants in the no only condition, $t < 1$, *ns*.

Table 2
Study 1: Means for the Primary Dependent Measures as a Function of Condition

Condition	RELAT	CORR	CONF
	Dependent measure		
Control	8.24_a	$.32_a$	7.59_a
Single explanation			
Negative only	5.87_a	$-.17_a$	7.39_a
No only	8.10_b	$.29_b$	7.97_a
Positive only	10.12_c	$.65_c$	8.27_a
Multiple explanation			
Negative–no	8.24_a	$.30_a$	8.20_a
Negative–positive	8.93_a	$.33_a$	7.70_a
No–negative	8.46_a	$.30_a$	8.00_a
No–positive	8.57_a	$.27_a$	8.29_a
Positive–negative	8.69_a	$.33_a$	7.83_a
Positive–no	8.38_a	$.26_a$	7.55_a

Note. RELAT = estimated degree of relationship between riskiness and success as a fire fighter (1 = *highly negative relationship,* 13 = *highly positive relationship*). CORR = estimated correlation between riskiness and success as a fire fighter. CONF = participants' confidence in their estimates (1 = *not at all confident,* 13 = *very confident*). Means not sharing a common subscript differ significantly at $p < .05$.

Multiple Explanation Conditions

Given that we replicated the explanation bias, we next examined the debiasing effects of a second explanation task. Comparisons of the estimates made in the multiple explanation conditions with those made in the single explanation conditions indeed demonstrated the debiasing effects of the second explanation task (see Table 2). In all cases, estimates made by participants in the multiple explanation conditions were significantly less biased in the negative direction than were those of participants in the negative only condition, and they were significantly less biased in the positive direction than were those of participants in the positive only condition. In all cases, estimates made by participants in the multiple explanation conditions did not differ from each other or from estimates made by control participants, all ts < 1, *ns*.

It is interesting to note that it appears that each multiple explanation condition was effective in debiasing participants' judgments. But were these conditions equally effective? Indeed, a one-way ANOVA comparing participants' estimates across the six multiple explanation conditions revealed no significant differences (both Fs < 1, *ns*). Moreover, analyses that examined whether order of explanation made a difference revealed that, in all cases, order did not affect estimates, all ts < 0.50, *ns*. Given that order of explanation did not matter, we performed an analysis in which we collapsed across order of explanation and compared the estimates in the no–negative, positive–negative, and positive–no conditions. This analysis tested a critical prediction of the split-the-difference view. If participants are merely splitting the difference in making their likelihood estimates, we would expect that participants explaining both the negative relationship and no relationship would provide more negative likelihood estimates than those explaining both the positive relationship and no relationship, with the participants explaining both positive and negative relationships falling in between these other two conditions. In contrast, both the uncertainty and the simulation heuristic explanations predict no differences in likelihood estimates across these conditions. This analysis revealed no differences across conditions on either the estimated relationship or the mean correlation coefficient measure (both $Fs < 1$, *ns*). Thus, it does not appear that our participants simply split-the-difference in making their likelihood estimates.

Additional analyses of the judgment data, as well as an examination of the confidence data, also allowed us to test critical predictions derived from the uncertainty argument. According to this view, participants in the multiple explanation conditions, confused and uncertain about the actual relationship between these variables, would be expected to (a) endorse responses near the middle (indifference point) of the judgment scale and (b) express lower confidence in their judgments than participants in the single explanation conditions. However, contrary to this explanation, participants' judgments in the multiple explanation conditions, like those of control participants, differed significantly from the midpoint of the judgment scale (and reflected a net positive relationship between riskiness and success as a fire fighter), $ts > 2.50$, $ps < .05$. In addition, an examination of the confidence data indicated no differences in confidence across conditions ($F < 1$, *ns*). Contrary to the uncertainty view, participants in the multiple explanation conditions were just as confident in their judgments as were participants in the single explanation conditions or the control condition. Thus, it does not appear that these effects can be attributed to uncertainty.

Analysis of Participants' Written Justifications of Their Judgments

Finally, we examined participants' open-ended responses to the question asking them to justify their judgments. These responses were coded by two independent raters according to two major dimensions. First, we noted the number of different possible relationships expressed between the variables of riskiness and fire fighting success. Answers that expressed only a single relationship (e.g., "Fire fighters must often take risky actions to save lives") were coded as 1, whereas answers that expressed two relationships that could exist between these variables (e.g., "One must be willing to take risks to be a successful fire fighter. On the other hand, being too risky can be bad because it could put one's own and others' lives in danger") were coded as 2. Some answers actually considered all three possible relationships and were coded as 3. Data from this measure could provide evidence of the extent to which participants were considering alternative relationships in making their judgments. Second, we examined the extent to which participants expressed uncertainty in their responses by the use of qualifiers (e.g., perhaps, maybe, or possibly) or hedging (e.g., "I don't know but ..." or "Then again it could be the case that ..."). Responses that included expressions of uncertainty were coded as 1, whereas responses stated confidently were coded as 0. Agreement for both ratings was very high (91% for number of relationships and 96% for hedging); disagreements were resolved by discussion.

ANOVAs were then performed on these data, comparing the responses of participants in the control (no explanation), single explanation, and multiple explanation conditions on each of the measures. An ANOVA on the hedging measure revealed no significant differences as a function of explanation condition, $F(2, 300) = 0.38$, *ns*. Counter to the uncertainty view, participants in the multiple explanation conditions did not express greater amounts of hedging and uncertainty in their responses (M = 27%) than participants in the single explanation conditions (M = 23%). However, an ANOVA on the number of relationships measure revealed a main effect of explanation condition, $F(2, 300) = 11.38$, $p < .001$. Participants in the multiple explanation ($M = 1.58$) and control ($M = 1.55$) conditions were more likely to mention additional relationships than were participants in the single explanation conditions ($M = 1.29$), both $ts > 3.10$, $ps < .01$. However, we further examined the different relationships expressed by participants in their responses to determine whether participants focused only on the explained relationships in their answers (as would be predicted by the split-the-difference view) or were just as likely to express relationships other than those previously explained (as predicted by the simulation heuristic view). Table 3 presents the number of participants who expressed positive, negative, or no relationships in their answers across all conditions. As Table 3 reveals, participants in each of the single explanation conditions were significantly more likely to express the previously explained relation-

Table 3
Study 1: Mean Number of Participants Mentioning a Given Relationship in Their Reasons for Judgment as a Function of Condition

Condition	Relationship: Positive	No	Negative	Total
Control	0.81_c	0.39_b	0.35_b	1.55
Single explanation				
Negative only	0.45_b	0.18_a	0.76_c	1.39
No only	0.47_b	0.57_{bc}	0.17_a	1.21
Positive only	1.00_d	0.16_a	0.09_a	1.25
Multiple explanation				
Negative–no	0.81_c	0.33_{ab}	0.33_{ab}	1.47
Negative–positive	0.78_c	0.56_b	0.22_a	1.56
No–negative	0.68_{bc}	0.60_b	0.36_b	1.64
No–positive	0.83_c	0.62_b	0.21_a	1.66
Positive–negative	0.79_c	0.50_b	0.32_{ab}	1.61
Positive–no	0.83_c	0.46_b	0.46_b	1.75

Note. Total refers to the mean total number of relationships expressed by participants in a given condition. This total number is broken down into the number of positive, negative, and no relationships between riskiness and success as fire fighter expressed by participants in that condition. Means not sharing a common subscript differ significantly at $p < .05$.

ship than either of the alternative relationships, all $Fs(2, 97) > 9.00$, $ps < .001$. For example, participants expressed a negative relationship more in the negative only condition ($M = .76$) than in the positive only or no only condition ($Ms = .09$ and $.17$, respectively). Yet, in the multiple explanation conditions, no such trend was observed; instead, participants were just as likely to discuss a given relationship when they had not previously explained that relationship as when they had (all $Fs[5, 148]/ < 1.10$, *ns*). These results argue strongly against the split-the-difference view and suggest that our participants were indeed considering alternatives in addition to those specified in the earlier explanation tasks, a finding consistent with the simulation heuristic view.

Discussion

The results of Study 1 suggest that explanation of a different relationship is enough to debias participants' likelihood judgments. Participants who explained no relationship between riskiness and fire fighting ability showed amounts of debiasing similar to those of participants who explained the opposite relationship (evidence counter to the predictions of the split-the-difference view). In all cases, the judgments of participants in the multiple explanation conditions significantly differed from the indifference point of the scale; furthermore, the confidence of these participants did not differ from that of participants in either the single explanation conditions or the control condition (evidence counter to the predictions of the uncertainty view). Thus, the results of Study 1 argue strongly against both the split-the-difference and uncertainty explanations for the effects of counterexplanation task.

Instead, the results of Study 1 are entirely consistent with the simulation heuristic view. The judgments of participants in the multiple explanation conditions did not differ from those of control (no explanation) participants, suggesting that the second explanation task completely removed the effects of the prior explanation task. Moreover, an analysis of participants' reasons for their judgments revealed that those in the multiple explanation conditions (like control participants) were more likely to express multiple alternative relationships between the variables of interest; furthermore, these relationships were just as likely to be ones that they were not explicitly asked to consider in the earlier explanation tasks as ones that they had previously explained. Thus, it appears that participants in the multiple explanation conditions were considering multiple alternative outcomes to the event in making their judgments.

However, several important issues remain unresolved by the results of Study 1. Specifically, Study 1 provides little in the way of direct evidence regarding the process by which the multiple explanation task debiases participants' judgments. Participants may indeed actively simulate alternative outcomes to the event; however, the multiple explanation task may also lead to debiasing as a result of simply debiasing the recall of the relevant information. Our earlier work (Hirt & Sherman, 1985; S. J. Sherman et al., 1983) showed that explanation tasks biased the recall of relevant evidence. Thus, it may be that counterexplanation tasks simply make accessible in memory a more unbiased set of evidence, resulting in unbiased judgments. This distinction highlights the difference between the availability and simulation heuristics noted by Kahneman and Tversky (1982). The availability heuristic refers to situations in which participants base their judgments on the information most readily accessible and retrieved from memory; thus, availability reflects the *retrieval* of previously stored information from memory. The simulation heuristic, in contrast, refers to situations in which participants base judgments on the ease with which they can generate scenarios consistent with a particular outcome; thus, simulation reflects an active *generation* process. It is possible that the debiasing effects of the multiple explanation task could be the result of either of these processes: Debiasing could be the product of unbiased retrieval of the relevant information (a recall-based process consistent with the availability heuristic) or the active simulation of multiple alternatives of the event (the simulation heuristic view). The simulation heuristic view makes no claims about the necessity of unbiased recall for debiasing, whereas the recall-based view does.

The results of Study 1 do not allow us to distinguish between these two processes. Although most of the participants in the multiple explanation conditions did not explain what would strictly be considered the "opposite" relationship, explanation of no relationship would still be likely to lead to retrieval of evidence suggestive of a different relationship. Thus, a recall-based argument could easily explain the obtained pattern of results. To evaluate the necessity of unbiased recall for debiasing, we needed to identify a counterexplanation task that would not produce unbiased recall. We believed that a critical test of the implications of these two processes would be to have participants perform a second explanation task in which they would explain a more extreme version of the same outcome. In this situation, participants are really not engaging in a counterexplanation task but are explaining an alternative version of the same outcome. Under such a condition, recall should be strongly biased in favor of the explained outcome; thus, a recall-based view would predict no debiasing. However, if such a condition leads to levels of debiasing similar to those found in counterexplanation conditions, there would be strong evidence in support of the simulation heuristic view that merely considering alternative outcomes breaks the inertia resulting from the conditional reference frame (Koehler, 1991) and leads participants to consider multiple alternatives for the target event. To address these issues, we performed a second study.

Study 2

Study 2 was designed to examine the limits of the debiasing effects of multiple explanation tasks. Study 2 used the same materials and procedure used in Hirt and Sherman (1985, Experiment 2). Participants were presented information concerning two mythical high school football teams, Norwood and Medway, about to play in an upcoming game. After reading detailed information about both teams, participants were asked to explain one of four outcomes: a convincing win by Norwood, a close win by Norwood, a close win by Medway, or a convincing win by Medway. Participants in the multiple explanation conditions were then asked to explain a second, alternative outcome. We included conditions in which participants made multiple explanations for the same outcome (e.g., a close win by Norwood and a convincing win by Norwood) to assess whether simply engaging in a multiple explanation task is sufficient to produce debiasing.

In addition, as in Hirt and Sherman (1985), Study 2 included recall measures for the presented information. Participants were asked to recall as much of the presented information as they could. After this free recall task, participants were asked to indicate, for each fact, which team that fact favored. In this way, we could obtain evidence for the bias inherent in participants' recall. In addition, we could examine the correlations between participants' recall bias and their likelihood judgments.

Method

Participants

Participants were 568 students enrolled in introductory-level psychology courses at the University of Wisconsin–Madison. Participants received course extra credit for their participation. Participants were tested in groups ranging in size from 2 to 25. Data from 17 participants were dropped because these participants failed to explain the proper event(s). Three additional participants were dropped because they failed to complete the dependent measures. Thus, data from 548 participants remained for inclusion in the subsequent analyses.

Procedure

The procedure used in Study 2 paralleled that used in Hirt and Sherman (1985, Experiment 2). On arrival, participants received a booklet containing all of the instructions and tasks that were relevant for their participation. Participants were initially told that the experiment was concerned with the ways in which people think about sporting events on the basis of information about the competing teams. Participants were then told that they would be asked to respond to information that they would be provided about football. However, participants were told that, before proceeding with these tasks, they would be asked a few questions about their background knowledge of football. They were then given a brief 10-item test designed to assess their general knowledge of contemporary college and professional football (cf. Hirt & Sherman, 1985).

On completion of the knowledge measure, participants were told that they would be reading a passage about two rival high school football teams, Norwood and Medway, that would be playing each other within the next few weeks. The passage contained detailed factual information about the teams, including team records; the scores of all games; state rankings; rushing, passing, and kicking statistics; defensive statistics; injury reports; and so forth. The passage had been used previously by Hirt and Sherman (1985, Experiment 2). In addition, given that this study was manipulating the magnitude of the outcome (small vs. large victory by one team or the other), we needed to ensure that participants viewed both a close game and a "blowout" as possible. Thus, included in the background data for the game was a quote from a local sportswriter who emphasized how emotional the game was and how difficult the outcome of such games is to call. The following quote was included: "In cross-town rivalries such as these, emotions run high, and you're just as likely to see a blowout for one team as you are to see a close game."

Participants were told that, once they had read the passage, they would be asked to recall facts about the teams without referring back to the passage. Thus, participants were given a recall set as they read the passage. Previous research by Hirt and Sherman (1985) and S. J. Sherman et al. (1983) showed that the explanation bias was particularly strong when participants were given a recall set at the time of encoding the passage information. Thus, the choice of a recall set allowed us to examine the debiasing effects of the multiple explanation conditions more clearly.

Participants were given 7 min to read over the passage. After reading the passage, experimental participants were presented with an explanation task. These participants were told to imagine that the game had been played and that Norwood (Medway) won the game by a margin of less than a touchdown (more than a touchdown). Thus, four different explanation tasks were included: Norwood convincing win (NOR++), Norwood close win (NOR+), Medway close win (MED+), and Medway convincing win (MED++). Participants were asked to write down those factors and reasons that might help one to explain a given outcome. As in Study 1, control participants completed a set of math problems in place of the explanation task to equate for time.

After the initial explanation task, participants in the multiple explanation conditions were reminded that their previous task was to imagine and explain a particular outcome of the game; however, they were reminded as well that the actual outcome of the (upcoming) game was yet to be determined. They were then asked to imagine and explain a different outcome of the game. For these multiple explanation participants, the following page included a different explanation task. All possible combinations of the multiple explanation conditions were included (a total of 12): NOR++/NOR+, NOR++/MED+, NOR++/MED++, NOR+/NOR++, NOR+/MED+, NOR+/MED++, MED+/NOR++, MED+/NOR+, MED+/MED++, MED++/NOR++, MED++/NOR+, and MED++/MED+. As a means of equating for time, participants in the single explanation and control conditions completed a set of math problems in place of the second explanation task.

After the explanation task(s), participants completed a series of judgment measures. These judgment measures assessed subjective likelihood estimates for the upcoming Norwood–Medway game. The primary judgment measures were the estimated probability of winning, ranging from *Norwood will definitely win* (1) to *Medway will definitely win* (13), and the predicted final score of the game. As in Study 1, participants also rated how confident they were in their predictions.

In addition to the judgment measures, participants completed a free-recall measure. This measure asked participants to list all of the information that they could remember from the passage and then to go back and rate, on a 4-point scale ranging from *not important* (1) to *very important* (4), the importance of each piece of information in making a prediction about the game's outcome. They were also asked to indicate which team each fact supported (Norwood, Medway, neither team, or both teams).

As in our previous work (Hirt & Sherman, 1985; S. J. Sherman et al., 1983), the order of completion of these dependent measures was counterbalanced; half of the participants filled out the judgment measures first and then recalled the information (judgment–recall order), and half recalled the information before making their judgments (recall–judgment order). In this way, we could examine the degree to which judgments affect or are affected by recall.

Results

Recall Data

We initially examined the effects of a multiple explanation task on recall. In our past work (Hirt & Sherman, 1985; S. J. Sherman et al., 1983), we found that both recall and judgments were strongly biased by the explanation task. Thus, we expected recall in all conditions to be strongly affected by the explanation task or tasks given. As in our previous work (Hirt & Sherman, 1985; S. J. Sherman et al., 1983), our primary recall index was a recall bias measure. This measure was computed for each participant from his or her free recall by calculating the ratio of facts favoring Medway to the sum of facts favoring Medway plus facts favoring Norwood (Medway/[Medway + Norwood]).

Table 4
Study 2: Means for the Combined Judgment Measure and Recall Bias Index as a Function of Condition

Condition	Dependent measure: Judgment	Recall bias
Control	$-.17_b$	$.624_b$
Single explanation		
NOR++	$-.58_a$	$.529_a$
NOR+	$-.25_{ab}$	$.568_{ab}$
MED+	$.12_{bc}$	$.614_b$
MED++	$.40_c$	$.713_c$
Multiple explanation		
NOR++/NOR+	$-.01_b$	$.557_a$
NOR++/MED+	$.05_b$	$.595_b$
NOR++/MED++	$.04_b$	$.618_b$
NOR+/MED+	$.09_{bc}$	$.643_b$
NOR+/MED++	$.05_b$	$.597_b$
MED+/MED++	$.10_{bc}$	$.686_c$

Note. Given that, in all cases, order did not affect results on any of the dependent measures, the means for the multiple explanation conditions are collapsed across order of explanation. The combined judgment index was computed by first transforming the probability of winning and point difference estimates into z scores and then averaging the two z scores. Higher numbers indicate greater support for a Medway victory. The recall bias index was computed by taking the number of facts favoring Medway and dividing by the number of facts favoring Medway plus the number of facts favoring Norwood (Medway/[Medway + Norwood]). Again, higher percentages indicate recall biased in favor of a Medway victory. Means not sharing a common subscript differ significantly at $p < .05$. NOR = Norwood; MED = Medway.

Facts favoring neither team were ignored.[5] Table 4 presents the recall bias data for all of the experimental conditions.

Single explanation conditions. Initially, we examined differences in recall bias in the single explanation conditions. A 4 (explanation condition) × 2 (order) ANOVA revealed a main effect of explanation condition, $F(3, 164) = 11.71$, $p < .001$. Participants in the MED++ condition showed significantly greater recall bias favoring Medway than participants in the MED+, NOR+, NOR++, or control condition, all $ts > 2.90$, $ps < .01$. In addition, participants in the MED+ condition showed greater recall bias favoring Medway than participants in either the NOR+ or NOR++ condition, both $ts > 2.25$, $ps < .03$. Participants in the NOR++ condition tended to show greater recall bias in favor of Norwood than did control participants, $t(85) = 2.77$, $p = .007$, and participants in the NOR+ condition, $t(164) = 1.80$, $p = .073$. Thus, replicating our previous work, recall was strongly biased in favor of the team explained.

Multiple explanation conditions. We next examined the recall bias in the multiple explanation conditions. Indeed, we predicted that recall in the multiple explanation conditions would vary as a function of the events explained. In the counterexplanation conditions (in which participants explained both a Norwood and a Medway victory), we expected that the multiple explanation task would make information favoring both teams accessible, resulting in unbiased recall. However, in the conditions in which participants explained two versions of the same event (either two Norwood victories or two Medway victories), we expected that recall would be strongly biased in the direction of the team explained. Thus, as with the single explanation conditions, we expected that the recall bias of participants in the NOR++/NOR+ condition should differ significantly from the recall bias of participants in the MED++/MED+ conditions and that the bias of participants in both of these conditions should differ significantly from that of participants in the control and counterexplanation conditions.

A Condition × Order ANOVA revealed a significant main effect of condition, $F(11, 319) = 3.13$, $p = .001$. As predicted, participants in the MED++/MED+ condition displayed greater recall bias favoring Medway than did participants in any other multiple explanation condition, all $ts > 2.50$, $ps < .01$. Likewise, participants in the NOR++/NOR+ condition displayed greater recall bias favoring Norwood than did participants in the other multiple explanation conditions, all $ts > 1.99$, $ps < .05$, except for the NOR++/MED+ condition, $t(338) = 1.08$, *ns*. Moreover, comparisons with the respective single explanation conditions provided further evidence of these differences in recall bias across multiple explanation conditions. Recall was significantly less biased in the MED++/NOR+ and MED++/NOR++ conditions than in the MED++ single explanation condition, both $ts > 3.20$, $ps < .005$; however, recall bias in the MED++ and MED++/MED+ conditions did not differ, $t(95) = 0.95$, *ns*. Likewise, recall in the NOR++/MED+ and NOR++/MED++ conditions was significantly less biased in favor of Norwood than was recall in the NOR++ single explanation condition, both $ts > 2.20$, $ps < .03$, but the NOR++ and NOR++/NOR+ conditions did not differ, $t(101) = 1.30$, *ns*. Indeed, only in the MED++/MED+ and NOR++/NOR+ conditions did recall bias differ significantly from that in the control condition, both $ts > 1.97$, $ps < .05$.[6]

[5] As in Hirt and Sherman (1985), we also computed a weighted recall bias measure for each participant by weighing all facts by participants' importance ratings. This measure was correlated with the recall bias measure ($r = .94$, $p < .0001$); moreover, the same pattern of results was obtained on both measures. We chose to report the unweighted recall bias mea-sure because past reviewers have argued that the weighted recall bias measure can be construed as another judgment measure rather than a pure measure of recall; thus, correlations between judgments and this weighted recall measure cannot be viewed as recall–judgment correlations.

We also performed analyses on the total number of facts recalled. Main effects of order, $F(1, 491) = 7.68$, $p = .006$, and of knowledge, $F(1, 491) = 27.66$, $p < .001$, were obtained on this measure. Participants in the recall–judgment order tended to recall more facts ($M = 13.1$) than did participants in the judgment–recall order ($M = 12.4$). To no surprise, high-knowledge participants recalled significantly more facts ($M = 13.2$) than did low-knowledge participants ($M = 11.1$). However, no main effects or interactions were obtained as a function of condition. Thus, any differences in recall bias cannot be attributed to differences in the total number recalled.

[6] We also obtained a significant Condition × Order of Dependent Measures interaction on the recall bias measure, $F(11, 319) = 2.08$, $p = .021$. Individual comparisons across order revealed that the effects of explanation condition on recall bias were more pronounced in the recall–judgment order than in the judgment–recall order. Recall bias in the MED++/MED+ condition favored Medway to a greater extent in the recall–judgment order than in the judgment–recall order, $t(54) = 2.01$, $p = .049$. Similarly, recall bias in the NOR++/NOR+ condition favored Norwood to a greater extent in the recall–judgment order than in the judgment–recall order, $t(56) = 1.85$, $p = .07$. Recall bias in the

Judgment Data

Given that recall was indeed biased in the direction of the event(s) explained, we next examined the judgment data. These data provide the critical test of the recall-based judgment versus the simulation heuristic views. If the recall-based view is correct, then there should be debiasing in counterexplanation conditions only; judgments in the NOR++/NOR+ and MED++/MED+ conditions, like judgments in the single explanation conditions, should be biased in the direction of the event(s) explained. However, if the simulation heuristic view is correct, participants in all multiple explanation conditions, including those in the NOR++/NOR+ and MED++/MED+ conditions, should show debiased judgments.

The two main judgment measures collected in this experiment were the probability of winning estimate and the predicted score of the game. From the predicted score, we computed a point difference index (cf. Hirt & Sherman, 1985; S. J. Sherman et al., 1983), subtracting the points scored by Norwood from the points scored by Medway. In this way, higher numbers on both indexes indicated judgments favoring Medway, whereas lower numbers on both favored Norwood. These two indexes were highly correlated ($r = .78$, $p < .0001$). Furthermore, the same pattern of results was obtained on each measure. Thus, for simplicity of presentation, we computed a single judgmental index by first transforming both measures to z scores and then calculating the average z score for the two measures. As was the case earlier, higher z scores on this measure indicate judgments favoring Medway. The subsequent analyses report the results obtained for this composite judgment measure. Table 4 presents the data for the combined judgment measure for all of the experimental conditions.

Single explanation conditions. A Condition (NOR++ vs. NOR+ vs. MED+ vs. MED++) × Order ANOVA performed on the composite judgment measure for the single explanation conditions revealed a strong main effect of condition, $F(3, 162) = 9.76$, $p < .001$. Participants in the MED++ and MED+ conditions estimated a Medway victory as significantly more likely than participants in both the NOR+ and NOR++ conditions, all ts > 2.00, ps < .05. Moreover, participants in the MED++ condition displayed significantly more favorability toward Medway than did control participants, $t(84) = 3.13$, $p = .002$; participants in the NOR++ condition displayed significantly more favorability toward Norwood than did control participants, $t(87) = -2.05$, $p < .05$. Thus, as with the recall data, judgments in the single explanation conditions were significantly biased in favor of the event explained. No significant differences were obtained in participants' confidence ratings, $F < 1$, *ns*.

Multiple explanation conditions. We next examined the debiasing effects of multiple explanation. As in Study 1, we first performed comparisons of the estimates made in the multiple explanation conditions with the respective single explanation conditions. These analyses revealed strong evidence of debiasing in all of the multiple explanation conditions. As expected, participants in the counterexplanation conditions were significantly less biased in favor of Medway than were participants in the MED++ condition, all ts > 1.96, ps < .05, and they were significantly less biased in favor of Norwood than were participants in the NOR++ condition, all ts > 2.00, ps < .05. However, more important, participants in the MED++/MED+ condition also were less biased in favor of Medway than were participants in the MED++ condition, $t(69) = 2.19$, $p = .032$. Likewise, participants in the NOR++/NOR+ condition showed less favoritism toward Norwood than did participants in the NOR++ condition, $t(72) = 3.22$, $p = .002$. Thus, it appears that participants in all multiple explanation conditions showed strong evidence of debiasing. Indeed, a one-way ANOVA comparing participants' judgments across the 12 multiple explanation conditions revealed no significant differences as a function of condition, $F(11, 332) = 1.03$, *ns*. In addition, judgments in all of the multiple explanation conditions did not differ from each other or from judgments in the control condition, all ts < 1.56, *ns*. Thus, these data suggest that biased recall is not a necessary condition for debiasing judgments. These results parallel the findings obtained by Anderson (1983a); in that work, Anderson found that perseverance effects were not contingent on biased recall for the original data.

Recall–Judgment Correlations

We also examined the patterns of recall–judgment correlations in Study 2. Consistent with our past work (Hirt & Sherman, 1985; S. J. Sherman et al., 1983), the overall recall–judgment correlations in the single explanation conditions were very high ($r = .49$, $p < .001$), reflecting memory-based judgments (cf. Hastie & Park, 1986) in these conditions. In fact, the average within-cell correlation between judgment and recall in the single explanation conditions was .36 (with a range from .21 to .72, all ps < .05). However, a different pattern of results emerged in the multiple explanation conditions. The corresponding overall recall–judgment correlation in the multiple explanation conditions ($r = .31$, $p < .001$) was significantly lower, $z = 2.19$, $p < .03$. Moreover, the average within-cell correlation in these multiple explanation conditions was .15, with a range from −.15 to .35. In the key multiple explanation conditions, namely the NOR++/NOR+ and MED++/MED+ conditions, the average within-cell correlation was −.11, indicating that judgments were independent of recall in these conditions. Thus, the recall–judgment correlations provide additional evidence that participants in the multiple explanation conditions were not basing their likelihood judgments on their recall of the relevant information at the time of judgment.

Discussion

The results of Study 2 suggest that the mere act of having participants engage in a second explanation task is sufficient to debias their likelihood judgments. Participants who explained a

other conditions did not differ significantly as a function of order, all ts < 1.15, *ns*. Moreover, an analysis of the condition main effect within order conditions revealed that the condition main effect was significant for the recall–judgment order, $F(11, 159) = 3.65$, $p < .001$, but not for the judgment–recall order, $F(11, 161) = 1.51$, $p = .133$. Thus, it appears that the effects of explanation on recall were moderated when recall was preceded by the judgment task.

different version of the same outcome showed amounts of debiasing similar to those of counterexplanation participants; moreover, judgments made in all of these multiple explanation conditions did not differ from those made in the control condition, indicating that the second explanation task completely eliminated the effects of the initial explanation task. These data are particularly striking in that, even in conditions in which the second explanation task was more extreme in the same direction than the first explanation task (i.e., MED+/MED++ and NOR+/NOR++), judgments were still debiased. Thus, it does not appear that explicit consideration of the opposite or even another distinct alternative is necessary to produce debiasing. These results are consistent with the simulation heuristic view and support Koehler's (1991) notion that simply breaking participants out of the inertia induced by the initial explanation task (via a second explanation task) is sufficient for debiasing.

In addition, the results of Study 2 are instructive regarding the role of memory in debiasing judgments. According to the recall-based view, the information accessible for recall determines judgment. Therefore, to the extent that a counterexplanation task debiases participants' recall of the relevant information, judgments should likewise be unbiased; however, if the counterexplanation task fails to result in unbiased recall, the recall-based view predicts biased judgments. Alternatively, the simulation heuristic view argues that counterexplanation tasks lead participants to consider multiple alternative scenarios or outcomes for the event, resulting in unbiased likelihood judgments. The simulation heuristic view argues that unbiased recall is not a necessary condition for debiasing to occur. As in our previous work (Hirt & Sherman, 1985; S. J. Sherman et al., 1983), recall was indeed strongly biased in favor of the outcome(s) explained in all conditions, including multiple explanation conditions. However, in the critical multiple explanation conditions (NOR++/NOR+ and MED++/MED+), judgments were unbiased, even though recall was profoundly biased in favor of the outcome explained. Thus, the results of Study 2 argue strongly against a recall-based view of the debiasing effects of a counterexplanation task and support Anderson's (1983a) contention that biased recall is not required for explanation and counterexplanation effects to occur.

Nonetheless, several questions remain unanswered by Studies 1 and 2. Although the results of these studies effectively rule out several alternative explanations for the debiasing effects of counterexplanation tasks, they provide little in the way of direct support for the simulation heuristic view. Arguably, much of the support derives from the lack of a difference in the effectiveness of the various multiple explanation conditions. Thus, to argue strongly that our results support the simulation heuristic view, we needed to provide more conclusive evidence that participants are actually considering additional alternative outcomes in making their judgments. The paradigms used in Studies 1 and 2, involving largely dichotomous outcomes, were limited in their ability to provide these sorts of data; instead, we needed to use a task involving several different discrete outcomes. In this situation, we could better assess whether participants in multiple explanation conditions are more likely to consider additional outcomes than participants in single explanation conditions.

In addition, we wanted to test a critical process assumption inherent to the simulation heuristic view. We have argued that successful completion of the second explanation task leads participants to realize that the outcome of the event is not as predictable as they had previously thought and motivates the simulation of multiple alternative outcomes. This argument assumes that participants can construct a viable explanation for the second outcome and deem that outcome to be plausible. Thus, we hypothesized that the effectiveness of a multiple explanation task in debiasing judgments depends on the *perceived plausibility* of the alternative outcome. Indeed, previous work on mental simulation (e.g., Kahneman & Tversky, 1982; Taylor & Schneider, 1989) suggests that such a plausibility assessment is critical in counterfactual reasoning; we believe that plausibility plays a similar role in prediction (cf. Gettys & Fisher, 1979). Explanation of a plausible alternative outcome should result in the successful generation of several possible scenarios that suggest an alternative outcome; this realization might also lead one to realize that multiple alternative outcomes are possible and need to be considered in making a likelihood judgment.[7] In contrast, the explanation of an implausible alternative is difficult and should result in the generation of few, if any, viable scenarios supporting that outcome; this realization may lead one to conclude that alternative outcomes are unlikely and should maintain (if not enhance) belief in the inevitability of the explained outcome. If successful debiasing is contingent on the perceived plausibility of the alternative explained, this would validate an important assumption of the simulation heuristic view. Moreover, it would establish a theoretically important limiting condition on the explanation effect and would also have practical implications for the quality of decision making. With these two purposes in mind, we conducted a third study.

Study 3

Study 3 was designed to (a) provide direct evidence that participants in multiple explanation conditions consider multiple alternative outcomes and (b) examine the moderating role of event plausibility in the effects of counterexplanation tasks. To accomplish these goals, we had Study 3 participants make estimates about the 1993 winner of major league baseball's National League (NL) East division. This choice was strategic for several reasons. First, unlike the two previous studies, the present study used a real-world domain that was important to our participants. Thus, Study 3 would add to the generalizability of our results to less artificial, real-world judgment tasks. Second, the NL East consisted of seven teams vying for the division championship. Thus, unlike in the previous studies, the counterexplanation tasks would involve explanation of a discrete alternative outcome (i.e., another team winning the division) rather than differences in magnitude (in margin of victory or degree of relationship). Third, the choice of division allowed us to select the Montreal Expos, a team that American college participants were less likely to be biased either toward or against, as the target of explanation; according to several sports

[7] Indeed, a recent study by Koehler (1994) suggests that explicitly asking participants to generate their own hypotheses before making predictions (as opposed to simply evaluating a prespecified set of hypotheses) led them to consider more alternative hypotheses, resulting in less overconfidence.

magazines, the Expos were considered the top candidate to win the division on the basis of their second-place finish the previous year. Finally, given that we wanted to include a range of multiple explanation tasks that varied in terms of their plausibility, the 1993 season afforded us a unique opportunity to vary the perceived plausibility of an alternative outcome: 1993 marked the first season of the Florida Marlins, an expansion team. Expansion teams are routinely awful in their initial seasons, making them clearly implausible contenders for a division championship. Moreover, the remaining contenders themselves differed in terms of their a priori likelihood of winning the division, allowing us to examine a broad range of alternatives that varied in their perceived plausibility.

All participants were presented with an initial set of information in the form of scouting reports of all seven NL East teams. After reading detailed information about all seven teams, experimental participants explained a division crown by the Montreal Expos. Participants in the multiple explanation conditions then explained an alternative outcome. Several alternative outcomes were used, including championships by the St. Louis Cardinals and Chicago Cubs (both reasonably plausible), as well as by the Florida Marlins (implausible). In addition, we included a condition in which participants could freely select a different team to explain from the remaining teams. It was predicted that explaining any alternative, so long as it was deemed plausible, would effectively counteract the effects of the initial explanation. A subjectively implausible alternative, however, should be completely ineffective in terms of debiasing participants. Thus, explaining the Cardinals, Cubs, or a different team of the participants' choosing should serve to debias participants' judgments, whereas explaining the Marlins should not.

The use of several alternative outcomes in Study 3 required two notable changes in the dependent measures. First, in the two previous experiments, participants made predictions along a continuum from outcome X to outcome Y, with an indifference point at the midpoint of the scale. In Study 3, we had participants make judgments along a scale reflecting the likelihood that the Expos would win the division. In this way, likelihood judgments indicated participants' belief in the focal hypothesis (cf. Koehler, 1991). However, this change afforded us the possibility to test more directly whether, consistent with the simulation heuristic view, explanation of a plausible alternative shakes participants' belief in the focal hypothesis, whereas explanation of an implausible alternative maintains their faith in the focal hypothesis. Second, the confidence measure used in Study 3 was also modified to reflect participants' confidence in their predictions about the Expos. Unlike the confidence measure used in the previous studies (in which participants merely indicated their confidence in their *judgment*), this measure assessed participants' confidence in the *focal hypothesis*. Thus, we would expect explanation of a plausible alternative to undermine participants' confidence in the focal hypothesis (relative to a single explanation condition), whereas explanation of an implausible alternative should maintain (or increase) their confidence in the focal hypothesis. Thus, debiasing would be revealed in Study 3 by reductions in both likelihood estimates and confidence estimates.

Finally, to obtain more direct evidence that participants are considering multiple alternatives in making their judgments, we included an additional dependent measure in which participants were asked to predict the final ranking of all seven NL East teams. The simulation heuristic view argues that explanation of a plausible alternative would lead participants to consider and simulate the prospects of additional teams (to those explained); thus, we would expect that participants in the multiple explanation conditions would realize that other teams were also viable contenders for the division crown, resulting in a decrease in the ranking of the Expos (relative to participants in the single explanation conditions) and a corresponding increase in the predicted ranking of the contender(s). Explanation of an implausible alternative should maintain these participants' faith in the focal hypothesis, leading them (like participants in the single explanation conditions) to ignore the prospects of other teams and rank the Expos highly.

Method

Participants

Individuals who considered themselves to be knowledgeable about and interested in baseball and baseball statistics were recruited for the experiment. Participants were 115 volunteers who were paid $5 to take part in the study. Participants were run in large groups, with all conditions represented in each group. The study was run during a 1-month period from May 20 to June 19, 1993, during the first half of the baseball season.

General Knowledge Measure

On arrival at the testing site, participants were told that they would be asked to respond to information that they would be provided about major league baseball. Participants were then given a series of 10 questions and items designed to assess their general knowledge of major league baseball (e.g., How far is the pitcher's mound from home plate? or Name last year's National League Cy Young Award winner). Scores on these items could range from 0 to 10.

Procedure

After completing the general knowledge measure, participants were told that they would be reading some information about the upcoming pennant race in the NL East, one of the (then) four divisions in major league baseball. Participants then received a packet containing detailed factual information about all of the teams in the NL East (i.e., Chicago Cubs, Florida Marlins, Montreal Expos, New York Mets, Philadelphia Phillies, Pittsburgh Pirates, and St. Louis Cardinals). This information included possible starting lineups for the 1993 season, probable pitching rotations, batting and pitching statistics from the previous year, and a short analysis of the strengths and weaknesses of each team.

Participants were given 15 min to read the information and were provided with recall set instructions similar to those used in Study 2. Participants were randomly assigned to one of six experimental conditions: (a) control, (b) explain Expos, (c) explain Expos–Cardinals, (d) explain Expos–Cubs, (e) explain Expos–Marlins, or (f) explain Expos–different team.

Control conditions. After reading the packet of information for 15 min, control participants were given the following instructions:

> Based on the information you read about the teams, we would like you to answer several questions about what you think *will* happen when the regular baseball season is actually over. Try to anticipate the future.

These participants then went on to the dependent measures.

Explain Expos condition. After reading the packet of information, participants in this condition were given the following instructions:

> One thing psychologists are interested in is how people explain hypothetical events. Of course, we don't know at the present time which team will win the pennant. However, we want you to *imagine* that the entire season has been played and that the *Montreal Expos* actually came in first place in the National League East. We are interested in what evidence, if any, you can write down which might help one to explain, or might have allowed one to anticipate a first place finish in the NL East by the *Montreal Expos.*

After writing their explanations, participants read the control condition instructions and then went on to the dependent measures.

Other conditions. Participants in the remaining conditions read the explain Expos instructions and wrote their explanations. They were then given the following instructions:

> Now we want you to *imagine* that the season has been played and that (the St. Louis Cardinals) (the Chicago Cubs) (the Florida Marlins) (a different team or teams) came in first place in the National League East. Once again, we are interested in what evidence, if any, you can write down which might help one to explain, or might have allowed one to anticipate (the St. Louis Cardinals) (the Chicago Cubs) (the Florida Marlins) (a different team or teams) coming in first place. Please approach this task as if you were doing it for the first time.

The "different team" instructions also included the following line: "Please choose a *specific* team or teams in the National League East." After writing their second explanations, participants read the control condition instructions and then went on to the dependent measures.

Dependent measures. The primary dependent measures were the estimated probability that the Montreal Expos would finish in first place in the NL East (ranging from 0% to 100%) and the predicted final rankings (finishes), made on a scale ranging from *first place* (1) to *last place* (7), of all of the teams in the division. In addition to providing data regarding the consideration of alternative outcomes, these latter rankings allowed us to check our intuitions regarding the perceived plausibility of each team finishing in first place. Participants also responded to a 14-point scale concerning how confident they were in their predictions of the Expos' season; the scale ranged from *not at all confident* (0) to *extremely confident* (13). Again, it is important to note that this confidence measure, unlike those used in the previous studies, measured participants' confidence in the *focal hypothesis.*

After completing these measures, participants were also asked whether they had any allegiance to, or a great deal of knowledge about, any particular team in the NL East or whether there was a team that they particularly disliked. Although approximately 10 participants were identified as being "unusually" biased for or against a particular team, inclusion or removal of these participants from the analyses did not affect the results in any significant way. Thus, the reported analyses included these participants. After completion of this final measure, participants were debriefed and thanked.

Results and Discussion

Predicted Finish

To examine the perceived plausibility of each team in the NL East coming in first, we initially examined the mean rankings for each team made by participants in the control condition. The rankings fell into three "levels" of plausibility: the Phillies (M = 2.60) and Expos (M = 2.90) were perceived as the most plausible teams and were ranked significantly higher than any of the remaining teams, all $ts > 2.50$, $ps < .01$. The Cardinals (M = 3.75), Cubs (M = 3.75), Pirates (M = 3.85), and Mets (M = 4.40) were all seen as fairly plausible, and rankings among these four teams did not differ significantly, all $ts < 1.20$, *ns.* Finally, the Marlins (M = 6.75), as predicted, were seen as the least plausible of all; they were ranked significantly below all of the other six teams, all $ts > 3.00$, $ps < .001$.

Explanation Bias

A one-way ANOVA (six explanation conditions) was then run on participants' predictions of the probability that the Expos would finish in first place. The overall test was significant, $F(5, 109) = 5.27$, $p < .001$. As depicted in Table 5, those who explained the Expos thought it more probable (M = 68.63) that the Expos would finish in first than did those in the control group (M = 45.85), $t(109) = -3.56$, $p = .001$, replicating the explanation bias effect.[8]

Debiasing and Plausibility

We predicted that only considering a plausible alternative would serve as an effective debiasing strategy. As can be seen in Table 5, considering the fairly plausible possibility that either the Cardinals (M = 45.26) or the Cubs (M = 51.42) might finish in first place significantly reduced the biasing effects of explaining the Expos, both $ts > 2.65$, $ps < .01$. In addition, explaining a different team of the participants' choosing (M = 45.05) also served as an effective debiasing strategy, $t(109) = 3.64$, $p < .001$.[9]

Of particular interest was the finding that explaining a first-place finish by the Florida Marlins, the new expansion team, did not debias participants. The mean in this condition (M = 64.47) was significantly greater than those in the other multiple explanation conditions, all $ts > 2.00$, $ps < .05$, and did not differ from those in the explain Expos condition, $t < 1$, *ns.* Thus, it appears that an alternative must be perceived as plausible if it is to effectively debias participants. In an effort to examine the relationship between plausibility and probability judgments more closely, we computed the correlation between partici-

[8] In Study 3, a significant negative correlation was found in the explain Expos condition between participants' scores on the general knowledge measure and their predictions of the Expos finishing first ($r = -.41$, $p < .05$), indicating that the biasing effects of explaining the Expos were constrained to some extent by participants' prior level of baseball knowledge. Likewise, Study 2 obtained a significant interaction between knowledge and condition in the single explanation conditions, $F(3, 154) = 4.18$, $p = .007$. High-knowledge participants were less influenced by the explanation task than were low-knowledge participants. These results are consistent with those of Hirt and Sherman (1985), who found that, overall, the judgments of knowledgeable participants were less affected by the explanation task than were those of naive participants.

[9] The most popular teams chosen by participants in this condition were the Cubs (n = 10) and Phillies (n = 8), followed by the Pirates (n = 2), Mets (n = 2), and Cardinals (n = 1). Also, 2 participants failed to choose a specific team, opting instead to explain a "hypothetical" team that had the characteristics of a first-place team.

Table 5
Study 3: Means for Perceived Likelihood and Confidence That the Expos Will Finish First as a Function of Condition

	Dependent measure	
Condition	Likelihood	Confidence
Control	45.85_a	8.06_a
Single explanation (Expos)	68.63_b	9.53_b
Multiple explanation		
Expos–Cardinals	45.26_a	7.76_a
Expos–Cubs	51.42_a	7.98_a
Expos–Marlins	64.47_b	9.24_b
Expos–different	45.05_a	7.45_a

Note. Numbers reported for likelihood are percentages out of 100, with higher numbers indicating a higher subjective probability that the Expos will finish in first place. Numbers reported for confidence are on a scale ranging from *not at all confident* (1) to *extremely confident* (13). Means not sharing a common subscript differ significantly at $p < .05$.

pants' rankings of the alternative they were asked to explain and their predictions of the Expos coming in first. Including only participants in the Cardinals, Cubs, Marlins, and different team conditions and partialing out the effects of the explanation condition to which they had been assigned, we found that the more implausible participants thought it was that the alternative team they were explaining would finish in first place, the higher their predictions of the Expos coming in first (i.e., the more biased they were in their predictions; $r = .43$, $p < .001$). This finding further highlights the important role of perceived plausibility in determining debiasing effectiveness.

The various explanation conditions also seemed to affect participants' confidence in their predictions concerning the Expos' finish (i.e., the focal hypothesis). A one-way ANOVA on participants' confidence ratings revealed a significant effect of explanation condition, $F(5, 109) = 2.42$, $p = .04$. As can be seen in Table 5, "biased" participants (i.e., those in the Expos [M = 9.53] and Marlins [M = 9.23] conditions) were more confident in their predictions than those who had been "debiased" (i.e., those in the Cardinals [M = 7.76], Cubs [M = 7.98], and different team [M = 7.45] conditions), $t(109) = 3.34$, $p = .001$. Thus, as predicted, consideration of a plausible alternative undermines belief in the focal hypothesis, resulting in lower confidence in that hypothesis; however, consideration of an implausible alternative not only biased subjective probability judgments but also made people feel more confident about their predictions concerning the Expos' finish.

Predicted Final Rankings of the Teams

Finally, we examined participants' predicted rankings of the various teams at the end of the season. Table 6 presents these data. We predicted that "biased" participants (i.e., those in the Expos and Marlins conditions) would predict a higher final ranking for the Expos than would "debiased" participants. A one-way ANOVA on these rank data revealed a significant main effect of explanation condition, $F(5, 109) = 3.33$, $p = .008$. Indeed, participants in the Expos (M = 1.63) and Marlins (M = 2.26) conditions assigned the Expos a higher final ranking than did participants in the Cardinals (M = 3.26), Cubs (M = 2.95), and different team (M = 3.00) conditions, $t(109) = 3.70$, $p < .001$.

However, further examination of these data provided an opportunity to gather more direct evidence that participants in the "debiasing" conditions were actually considering multiple alternative outcomes. We first examined the extent to which participants in the Cardinals, Cubs, and Marlins conditions elevated their final rankings of their (respective) explained teams. Interestingly, as can be seen in Table 6, we found that in none of these conditions did the mean rank given to the explained team significantly increase, both $ts < 1$, *ns*. Instead, however, we found that, in all of the debiasing conditions, participants' mean ranking of the Phillies improved. It is important to note that, during the 1993 season, the Philadelphia Phillies got off to a surprisingly fast start and ended up winning the NL East pennant. Moreover, at the time of the study, knowledgeable participants were aware that the Phillies were leading the division. Thus, at the time, the Phillies were the most salient alternative to the Expos. An ANOVA performed on participants' mean final ranking for the Phillies revealed a significant main effect of explanation condition, $F(5, 109) = 2.31$, $p = .05$. Participants in the debiasing conditions, namely the Cardinals (M = 2.00), Cubs (M = 1.95), and different team (M = 2.00) conditions, ranked the Phillies significantly higher than did participants in the biased conditions, namely the Expos (M = 3.16) and Marlins (M = 3.26) conditions, $t(109) = 3.38$, $p = .001$.

Another way to look at these data would be to compare the ranks given to the Expos as opposed to the Phillies in the respective explanation conditions. Recall that, in the control condition, participants rated the Phillies and Expos as equally plausible winners of the division. Indeed, an examination of the relative ranks revealed that participants in the debiased conditions ranked the Phillies significantly higher than the Expos, whereas participants in the biased conditions ranked the Expos significantly above the Phillies. On the basis of these observations, we conducted a repeated measures ANOVA with explanation condition as the between-subjects variable and team (Expos vs. Phillies) as the repeated measures variable. This analysis revealed a significant Explanation Condition × Team interaction, $F(5, 109) = 4.34$, $p = .001$. Thus, these data provide clear evidence that participants given the task of explaining a plausible alternative were more likely to consider at least one additional alternative to the one explained, resulting in qualitatively different judgments of the future outcome. Conversely, these data indicate that participants asked to explain an implausible alternative, like participants in the single explanation conditions, failed to consider salient alternative outcomes, suggesting that such an explanation task fails to break the inertia caused by the initial explanation task.

General Discussion

The results of the present research provide strong evidence for the debiasing effects of a counterexplanation task. Participants asked to explain an alternative outcome, even an alternative version of the same outcome, exemplified judgments that did not differ from those of control (no explanation) participants, indicating that the counterexplanation task completely

Table 6
Study 3: Mean Final Ranking Assigned to Each Team as a Function of Condition

Condition	Dependent measure				
	Expos	Phillies	Cardinals	Cubs	Marlins
Control	2.90_a	2.60_{ab}	3.75_c	3.75_c	6.75_d
Single explanation (Expos)	1.63_b	3.16_a	3.58_{ac}	3.21_{ac}	6.63_d
Multiple explanation					
Expos–Cardinals	3.26_{ac}	2.00_b	3.26_a	4.16_c	6.63_d
Expos–Cubs	2.95_a	1.95_b	3.84_c	3.68_c	6.58_d
Expos–Marlins	2.26_b	3.26_a	4.42_c	3.32_{ac}	6.89_d
Expos–different	3.00_{ac}	2.00_b	3.58_{ac}	3.63_{ac}	6.32_d

Note. Mean ranks could range from 1 (*first place*) to 7 (*last place*). Means not sharing a common subscript differ significantly at $p < .05$.

eliminated the effects of the prior explanation task on likelihood judgments. However, our results also identify the perceived plausibility of the alternative outcome as an important constraint on the effectiveness of counterexplanation tasks. Participants asked to explain subjectively plausible alternative outcomes showed debiasing, whereas those asked to explain subjectively implausible alternative outcomes maintained their belief in the focal hypothesis. Moreover, the greater the perceived plausibility of the alternative, the greater the debiasing effect of the counterexplanation task.

Specifying the Process by Which Counterexplanation Tasks Work

In this research, our primary interest was to specify the process by which counterexplanation tasks operate. In the introduction, we discussed several viable explanations for the effectiveness of counterexplanation tasks. However, we advocated a simulation heuristic explanation that proposed that the act of engaging in a counterexplanation task encourages participants to consider multiple alternative outcomes for an event in addition to those specified in the counterexplanation task. Indeed, the results of Study 3 provide convincing evidence in support of this argument. Even though they had not been asked to explain a Phillies pennant victory, participants in the multiple explanation conditions asked to explain a plausible alternative were more likely to pick the Phillies to win the NL East than participants were in the single explanation conditions. Thus, these multiple explanation participants were considering the likelihood of another salient alternative in addition to the likelihood of the two teams explained. On the other hand, multiple explanation participants who explained an implausible alternative, like single explanation participants, continued to believe that the Expos would win the division. Thus, these participants remained fixated on the likelihood of the initially explained outcome and failed to consider the salient alternative outcome in their likelihood judgments.

Our results also allow us to successfully rule out several alternative hypotheses regarding process. For instance, the results of Studies 1 and 2 effectively ruled out both split-the-difference and uncertainty explanations for these debiasing effects. Furthermore, the results of Study 2 also clarify the role of recall in the effectiveness of counterexplanation tasks. According to a recall-based view, counterexplanation tasks are effective to the extent that they activate the recall of additional information that may question the likelihood of the initially explained outcome or imply an alternative outcome. The results of Study 2 indicate that unbiased recall is not a necessary condition for debiasing. Participants in the NOR++/NOR+ and MED++/MED+ conditions evidenced recall strongly biased by the explanation task but still showed unbiased judgments. The recall–judgment correlations in these conditions were nearly zero, implying that judgments were independent of the information recalled. Thus, our results argue strongly against a recall-based view of the debiasing effects of counterexplanation tasks.

Instead, we believe that the effects of counterexplanation tasks are best understood in terms of the operation of the simulation heuristic (Kahneman & Tversky, 1982). Our earlier work (Hirt & Sherman, 1985; S. J. Sherman et al., 1983) demonstrated that an explanation task leads participants to generate scenarios that result in the specified outcome. However, once participants are able to successfully generate a plausible explanation supporting the assigned outcome, they tend to cease the simulation process and fail to consider alternative scenarios that may imply a different outcome (cf. Shaklee & Fischhoff, 1982). A counterexplanation task requires participants to generate scenarios that result in an alternative outcome. To the extent that participants are able to successfully generate several plausible scenarios implying this alternative outcome, their confidence in the focal hypothesis is undermined. In addition, the present results imply that the successful simulation of an alternative outcome resulting from the counterexplanation task prompts participants to consider additional alternative outcomes. As a result, it appears that the counterexplanation task breaks participants' inertia with regard to the focal hypothesis and leads to a more thorough and comprehensive consideration of the likely outcome of the event.

Motivational Versus Cognitive Mechanisms Underlying Debiasing

However, we acknowledge that the mechanisms that initiate this further consideration of alternative outcomes remain to be determined. Clearly, one might expect that, after generating a plausible scenario for the counterexplanation task, participants would cease further consideration of alternative outcomes, maintaining a truncated search strategy (cf. Shaklee & Fischhoff, 1982). In their research, Kruglanski and his colleagues (Kruglanski, 1990; Kruglanski & Mayseless, 1988; Mayseless & Kruglanski, 1987) have argued that efforts to debias or "unfreeze" participants require the arousal of epistemic motives. For instance, Mayseless and Kruglanski (1987) found that participants under high fear of invalidity conditions generated and considered more hypotheses concerning an object's identity in a tachistoscopic recognition task. Thus, the counterexplanation task may enhance participants' fear of invalidity and increase their motivation to draw accurate judgments. Indeed, the counterexplanation task may have suggested to participants that the prediction task is not so straightforward and may have represented a challenge to prognosticate the likely outcome. Moreover, this motivational view can also explain why participants asked to consider an implausible alternative (e.g., the Marlins) fail to consider additional alternatives: Explanation of an implausible alternative may allay participants' fear of invalidity and increase their faith in the likelihood of the initially explained outcome.

The present argument highlights an important issue in the debiasing literature: Are these errors in judgment due to participants' lack of motivation (i.e., motivationally based) or limited cognitive capacities (cognitively based)? In a recent review of the debiasing literature, Arkes (1991) found that accuracy incentives do not improve judgmental accuracy. For example, Lord et al. (1984) found that warning participants about possible bias and instructing them to "be unbiased" does not reduce bias in their judgments; instead, these researchers found that participants showed debiased judgments only when they were given a counterexplanation task that explicitly instructed them to consider alternatives, suggesting that these errors are cognitively as opposed to motivationally based. However, a recent article by Schuette and Fazio (1995) provides evidence that increasing accuracy motivation can successfully debias judgments. Following the Lord, Ross, and Lepper (1979) paradigm, participants given a fear of invalidity manipulation (having their judgments compared with those of a panel of experts and also anticipating having to later justify their judgments to other participants) showed no biased assimilation of evidence, suggesting that enhancing participants' motivation to be accurate can eliminate this judgmental bias. This finding, along with those of Kruglanski and his colleagues, implies that these biases may have a strong motivational component.

Thus, an important question that needs to be addressed in future research is the extent to which these debiasing effects occur through motivational as opposed to cognitive mechanisms. Is increased motivation enough to encourage participants to consider and simulate additional alternatives, or do participants also need explicit prompting (through a counterexplanation task) to engage the simulation process? If, indeed, a counterexplanation task is necessary to "prime the pump" and engage the simulation process, important questions still remain about the features of such a task. Although our results indicate that explanation of any alternative outcome (even an alternative version of the same outcome) is sufficient to produce debiasing, it is important to note that, in all cases, participants were explaining an alternative outcome of the same event; thus, they had to "undo" their earlier explanation and generate a second, different explanation. A more stringent test of the simulation heuristic view would be to have participants create counterexplanations that are unrelated to the original explanation (e.g., explain a positive relationship between introversion and fire fighting success or explain the New York Yankees winning the American League East crown) and, therefore, do not require any undoing.[10] If merely the act of engaging in a counterexplanation task is sufficient to engage the simulation process, one would expect these conditions to produce the same debiasing effect, provided that the explained outcome is plausible. Explanation of any implausible outcome should maintain participants' faith in the focal hypothesis and result in no debiasing. Although these questions await further research, they highlight the importance of considering both the motivational and cognitive underpinnings of judgmental biases.

[10] We thank an anonymous reviewer for making this point.

References

Anderson, C. A. (1982). Inoculation and counter-explanation: Debiasing techniques in the perseverance of social theories. *Social Cognition, 1,* 126–139.

Anderson, C. A. (1983a). Abstract and concrete data in the perseverance of social theories: When weak data lead to unshakeable beliefs. *Journal of Experimental Social Psychology, 19,* 93–108.

Anderson, C. A. (1983b). Imagination and expectation: The effect of imagining behavioral scripts on personal intentions. *Journal of Personality and Social Psychology, 45,* 293–305.

Anderson, C. A., & Godfrey, S. (1987). Thoughts about actions: The effects of specificity and availability of imagined behavioral scripts on expectations about oneself and others. *Social Cognition, 5,* 238–258.

Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology, 39,* 1037–1049.

Anderson, C. A., New, B. L., & Speer, J. R. (1985). Argument availability as a mediator of social theory perseverance. *Social Cognition, 3,* 235–249.

Anderson, C. A., & Sechler, E. S. (1986). Effects of explanation and counterexplanation on the development and use of social theories. *Journal of Personality and Social Psychology, 50,* 24–34.

Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin, 110,* 486–498.

Campbell, J. D., & Fairey, P. J. (1985). Effects of self-esteem, hypothetical explanations, and verbalizations of expectancies on future performance. *Journal of Personality and Social Psychology, 48,* 1097–1111.

Carroll, J. S. (1978). The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. *Journal of Experimental Social Psychology, 14,* 88–96.

Devine, P. G., Hirt, E. R., & Gehrke, E. M. (1990). Diagnostic and confirmation strategies in trait hypothesis testing. *Journal of Personality and Social Psychology, 58,* 952–963.

Fischhoff, B. (1975). Hindsight does not equal foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 288–299.

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422–444). Cambridge, England: Cambridge University Press.

Fischhoff, B., & Beyth, R. (1975). "I knew it would happen": Remem-

bered probabilities of once-future things. *Organizational Behavior and Human Performance, 13,* 1–16.
Gettys, C. F., & Fisher, S. D. (1979). Hypothesis plausibility and hypothesis generation. *Organizational Behavior and Human Performance, 24,* 93–110.
Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46,* 107–119.
Gregory, W. L., Cialdini, R. B., & Carpenter, K. M. (1982). Self-relevant scenarios as mediators of likelihood estimates and compliance: Does imagining make it so? *Journal of Personality and Social Psychology, 43,* 89–99.
Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review, 93,* 258–268.
Hirt, E. R., & Sherman, S. J. (1985). The role of prior knowledge in explaining hypothetical events. *Journal of Experimental Social Psychology, 21,* 519–543.
Hoch, S. J. (1985). Counterfactual reasoning and accuracy in predicting personal events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 719–731.
Johnson, M. K., & Sherman, S. J. (1990). Constructing and reconstructing the past and future in the present. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 482–526). New York: Guilford Press.
Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93,* 136–153.
Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). Cambridge, England: Cambridge University Press.
Kaplan, K. J. (1972). On the ambivalence-indifference problem in attitude theory and measurement: A suggested modification of the semantic differential technique. *Psychological Bulletin, 77,* 361–372.
Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin, 110,* 499–519.
Koehler, D. J. (1994). Hypothesis generation and confidence in judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 461–469.
Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 107–118.
Kruglanski, A. W. (1990). Motivations for judging and knowing: Implications for causal attribution. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 333–368). New York: Guilford Press.
Kruglanski, A. W., & Mayseless, O. (1988). Contextual effects in hypothesis testing: The role of competing alternatives and epistemic motivations. *Social Cognition, 6,* 1–21.
Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). Cambridge, England: Cambridge University Press.
Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology, 47,* 1231–1243.
Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37,* 2098–2109.
Mayseless, O., & Kruglanski, A. W. (1987). What makes you so sure?: Effects of epistemic motivations on judgmental confidence. *Organizational Behavior and Human Decision Processes, 39,* 162–183.
Ross, L., Lepper, M. R., Strack, F., & Steinmetz, J. (1977). Social explanation and social expectation: Effects of real and hypothetical explanations on subjective likelihood. *Journal of Personality and Social Psychology, 35,* 817–829.
Schuette, R., & Fazio, R. H. (1995). Attitude accessibility and motivation as determinants of biased processing: A test of the MODE model. *Personality and Social Psychology Bulletin, 21,* 704–710.
Shaklee, H., & Fischhoff, B. (1982). Strategies of information search in causal analysis. *Memory & Cognition, 10,* 520–530.
Sherman, R. T., & Anderson, C. A. (1987). Decreasing premature termination from psychotherapy. *Journal of Social and Clinical Psychology, 5,* 298–312.
Sherman, S. J., Cialdini, R. B., Schwartzman, D. F., & Reynolds, K. D. (1985). Imagining can heighten or lower the perceived likelihood of contracting a disease: The mediating effect of ease of imagery. *Personality and Social Psychology Bulletin, 11,* 118–127.
Sherman, S. J., Skov, R. B., Hervitz, E. F., & Stock, C. B. (1981). The effects of explaining hypothetical future events: From possibility to probability to actuality and beyond. *Journal of Experimental Social Psychology, 17,* 142–158.
Sherman, S. J., Zehner, K. S., Johnson, J., & 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.
Skov, R. B., & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology, 22,* 93–121.
Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 544–551.
Snyder, M., & Swann, W. B., Jr. (1978). Hypothesis-testing in social interaction. *Journal of Personality and Social Psychology, 36,* 1202–1212.
Taylor, S. E., & Schneider, S. K. (1989). Coping and the simulation of events. *Social Cognition, 7,* 174–194.
Thompson, M. M., Zanna, M. P., & Griffin, D. W. (in press). Let's not be indifferent about (attitudinal) ambivalence. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences.* Hillsdale, NJ: Erlbaum.
Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207–232.
Vallone, R., Griffin, D. W., Lin, S., & Ross, L. (1990). Overconfident prediction of future actions and outcomes by self and others. *Journal of Personality and Social Psychology, 58,* 582–592.
Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12,* 129–140.
Wason, P. C., & Golding, E. (1974). The language of inconsistency. *British Journal of Psychology, 65,* 537–546.
Wells, G. L., & Gavanski, I. (1989). Mental simulation of causality. *Journal of Personality and Social Psychology, 56,* 161–169.

Received October 10, 1994
Revision received July 12, 1995
Accepted July 14, 1995 ■