MEMORY, 2003, *11* (4/5), 421–441

Surprise, defence, or making sense: What removes hindsight bias?

Mark V. Pezzo

University of South Florida, USA

This paper examines predictions concerning the *absence* of hindsight bias. Some hypothesise that because hindsight bias increases with outcome ''surprisingness'', only unsurprising outcomes will remove it. Others suggest the opposite—that very surprising outcomes will reduce or reverse the bias. A proposed sense-making model suggests that unexpected outcomes (i.e., *initially* surprising) invoke greater sensemaking, which typically produces greater hindsight bias. If the process is not successful, however, the bias may be reduced or reversed. Expected outcomes will also produce little hindsight bias, but only because they invoke relatively little sensemaking in the first place. Feelings of surprise arising from sensemaking (i.e., *resultant* surprise) should be inversely related to hindsight bias. Results of four experiments provide support for the model. A secondary goal was to determine the boundaries of a defensive-processing mechanism also thought to reduce hindsight bias for negative, self-relevant outcomes. Results suggest that a sense of responsibility for the outcome may be necessary for defensive processing to be activated.

> ''I was surprised, but not that surprised. I mean, it makes sense.''
>
> —from an internet chat.

When musing over past events, how inevitable do they seem? A sizeable literature on the hindsight bias tells us that we are disposed to believe we ''knew it all along'' that such events would occur. In his seminal article, Fischhoff (1975) argued that this inclination is caused by a relatively automatic and unconscious *sense-making* process (creeping determinism) that focuses attention on outcome-consistent information and away from outcome-inconsistent information. More than two decades of research, and over 120 published articles[1] provide a wealth of evidence for the ubiquity of the hindsight bias—it has been shown in medical diagnoses, presidential elections, legal decisions, accounting, sporting events, and myriad other domains (see Christensen-Szalanski & Willham, 1991, and Hawkins & Hastie, 1990, for reviews). This paper, however, addresses those instances in which the hindsight bias does *not* occur.

Consider a stockbroker who advises her clients, in good faith, to purchase many shares of a certain stock, but then learns that the stock has ''crashed''. Will the stockbroker later believe that she knew that this would happen? Or did voters supporting President Clinton respond that they ''knew it all along'' that he was having an affair with Monica Lewinsky? Surely there are times when we don't think that we knew it all along. With few exceptions, however, the extant literature has not addressed the conditions under which we do not exhibit the hindsight bias.

[1] Christensen-Szalanski and Willham (1991) found 40 articles, reporting 128 experiments on hindsight bias. A recent search revealed over 120 articles, a likely underestimation, as articles can use terms besides ''hindsight bias'' (e.g., Alicke, Davis, & Pezzo, 1994; Tan & Lipe, 1997).

Requests for reprints should be sent to Mark V. Pezzo, Department of Psychology, University of South Florida, 140 7th Ave., South, St. Petersburg, FL 33701, USA. Email: pezzo@stpt.usf.edu

I thank Stephanie Parks, Hal Arkes, Mark Leary, Eric Stone, Emily Page, Maria Watson, William Fleeson, John Seta, and three anonymous reviewers for their helpful comments and suggestions.

I also wish to thank the many lab assistants who helped conduct this research: Adrianne Malanos, Jill Rosa, Peter Rives, Emily Greenwood, Matt Morris, Andrew Adamsbaum, Neal Wildrick, Karen Manship, David James, Laurie Best, Bethany Wolf, and Ellen J. Godfrey.

http://www.tandf.co.uk/journals/pp/09658211.html DOI:10.1080/09658210244000603

Some researchers have effectively reduced the hindsight bias by forcing participants to consider how alternative outcomes might have occurred (Arkes, Faust, Guilmette, & Hart, 1988; Davies, 1987; Nario & Branscombe, 1995; Slovic & Fischhoff, 1977). Although this approach is effective, it gives no indication of when people might *spontaneously* avoid the bias, either by ignoring the outcome altogether, or by considering information that is incongruent with it.

THE ROLE OF SELF-RELEVANT OUTCOMES

Mark and Mellor (1991) suggest that outcome self-relevance might play an important role in determining whether or not the hindsight bias occurs. They found that people who lost their jobs following a company layoff reported the layoff as less foreseeable than did either community members or employees who did not lose their jobs. Mark and Mellor held that only the laid-off workers did not exhibit the bias because they were motivated to reduce their own sense of culpability for not having acted to circumvent their layoff. However, their study provides only indirect evidence for this because they included no pre-outcome control condition. Participants were asked how ''foreseeable'' the outcome was only *after* the layoff had taken place. If laid-off workers felt the outcome was not particularly foreseeable *before* the layoff as well, then no hindsight bias would be said to exist. A regression-discontinuity technique was used to help reduce this possibility by statistically controlling for a number of demographic factors among the different groups, but the inclusion of a no-outcome control condition would provide more compelling evidence for such a claim.

A few researchers have used the no-outcome control design to examine the defensive-processing mechanism proposed by Mark and Mellor (1991). Pezzo (1996) hypothesised that highly committed decision makers would exhibit less hindsight bias if the outcome contradicted their decision. In two separate experiments, however, he found equivocal evidence for such an effect. Louie (1999), however, found that business students who believed that their stock choice had failed exhibited no hindsight bias, but students whose stock succeeded did exhibit the bias. Similarly, Louie, Curren, and Harich (2000) found that MBA students showed hindsight bias when the outcome was favourable to their team's decision, but not when it was unfavourable.

An interesting question for the defensive-processing hypothesis is whether a sense of culpability or responsibility for the negative outcome is necessary to produce the effect. Such a requirement is implied by the work of Mark and Mellor and others (e.g., Louie et al., 2000; Markman & Tetlock, 2000), but has never been directly tested. This is important, because there exist negative outcomes that can be both upsetting and relevant to one's self-esteem, but for which one would not feel culpable (Cialdini, Borden, Thorne, Walker, Freeman, & Sloan, 1976). Would such an outcome be sufficient to activate the defensive-processing mechanism?

THE ROLE OF SURPRISE

Perhaps more important for the defensive-processing approach than the issue of culpability is the fact that, in addition to being threatening, negative outcomes are often more *surprising* than positive outcomes (Taylor & Brown, 1988). Additionally, there is reason to believe that a self-relevant event, such as the layoff in Mark and Mellor's (1991) study, may be more surprising than one that is not self-relevant. For example, Falk (1989) found that coincidences with high personal significance were judged more surprising than those with low personal significance, which were judged less likely than coincidences that happen to other people. There is evidence to suggest that this effect is robust (cf. Zakay, 1984), and thus any examination of self-relevance should probably also take into consideration perceptions of outcome surprisingness.

What is the effect of surprise on the hindsight bias? Christensen-Szalanski and Willham (1991) note that ''nearly all [hindsight] researchers … have disregarded the possible moderating effect of an event's surprisingness on the hindsight bias, and the few studies that do attend to an event's surprisingness do not present the data that are needed to evaluate the hypothesized moderating influence'' (p. 152). In the decade since that article, however, a few researchers have made explicit predictions concerning the effect of surprise.

Mazursky and Ofir (1990, 1996) suggested that an extremely surprising outcome, alone, could reduce or even reverse the hindsight bias. They theorised that a surprising outcome triggers ''special processing'' which, along with feelings of

surprise, can cancel or reverse the effects of hindsight bias. They suggest that this processing involves a greater number of attempts at explanation, recall, and justification, which may offset the typical effects of creeping determinism. Note also that, according to their model, moderately or completely unsurprising outcomes should produce typical hindsight biases. In three separate experiments, using evaluations of medical decisions, consumer products, and modern art, they showed that people receiving extremely surprising outcomes gave hindsight estimates that were in the ''reverse'' direction (Ofir & Mazursky, 1997). That is, likelihood estimates for highly expected outcomes that did *not* occur were even *larger* in hindsight. Reverse hindsight bias is somewhat controversial, because previous claims (Guerin, 1982; Mazursky & Ofir, 1990; Verplanken & Pieters, 1988) have been criticised for misinterpretation of the data (Arkes, 1988; Hawkins & Hastie, 1990; Mark & Mellor, 1994). Still, others have reported such effects (Choi & Nisbett, 2000; Haslam & Jayasinghe, 1995; Menec & Weiner, 2000; Pezzo, 1996; Winman, 1997).

Despite the intuitive appeal of reverse hindsight bias, others have argued that surprising outcomes should produce *greater* rather than less hindsight bias (cf. Roese & Olson, 1996). As Mazursky and Ofir (1996) have noted, considerable evidence exists indicating that unexpected outcomes produce the greatest search for causal antecedents (Hastie, 1984; Pyszczynski & Greenberg, 1981; Sanna & Turley, 1996; Weiner, 1985). But this type of ''sense-making'' activity has been argued, by most, to produce hindsight bias rather than remove it (Fischhoff, 1975; Roese & Maniar, 1997; Schkade & Kilbourne, 1991; Wasserman, Lempert, & Hastie, 1991). From this logic it follows that outcomes that are most unexpected should produce the greatest hindsight bias.

Conversely, outcomes that are congruent with expectations should produce little hindsight bias because they require little or no search for causal antecedents. This prediction has received some empirical support; for example, Cannon and Quinsey (1995, Study 1) found that the occurrence of an outcome that was generally expected (estimated likelihood = 70%) did not produce hindsight bias, but its unexpected *non*-occurrence (estimated likelihood = 30%) did produce a significant hindsight bias.

Some indirect evidence indicating a positive relationship between surprise and hindsight magnitude can be gleaned from studies using almanac trivia questions. People typically exhibit greater hindsight bias when they are presented with difficult or misleading questions, which are thought to be more surprising than easy questions (Christensen-Szalanski & Willham, 1991; Fischhoff, 1977; Hoch & Loewenstein 1989; Winman, 1997).[2] However, the best example of this relationship comes from Schkade and Kilbourne (1991). Using business scenarios, they created high and low expectations for a positive outcome by manipulating performance history and employee behaviour. A surprising outcome was defined as incongruent with a priori expectations. Hindsight bias was relatively large for incongruent outcomes, but was smaller and sometimes eliminated for congruent outcomes. These findings are consistent with the idea that unexpected outcomes produce greater sense-making activity (Weiner, 1985). In response, Ofir and Mazursky (1997) suggest that Schkade and Kilbourne may not have used outcomes that reached the threshold of surprise sufficient to offset or reverse the bias. Whether or not this is the case, it is notable that Ofir and Mazursky *have* demonstrated reductions and even reversals of the bias.

A SENSE-MAKING MODEL OF HINDSIGHT BIAS

One possible solution requires that we consider two meanings for the term ''surprise''. The most common usage defines a surprising outcome as one that is incongruent with a priori expectations (Fischhoff & Beyth, 1975; Schkade & Kilbourne, 1991). Such surprise could be thought of as ''initial'' surprise, and it is this type of surprise that triggers sense-making activity (Hastie, 1984; Pyszczynski & Greenberg, 1981; Weiner, 1985). The second meaning for surprise refers to the phenomenological ''feeling'' that results from the sense-making process. Some initially surprising outcomes are surely easier to make sense of than others. Those that are more difficult, however, likely result in conscious awareness of the incongruity, and it is this awareness that we may call ''resultant'' surprise. Thus, an originally incon-

[2] Hoch and Loewenstein (1989) have been cited as evidence for a reverse hindsight bias with difficult items (e.g., Hawkins & Hastie, 1990; Louie, 1999; Louie et al., 2000; Pohl, 1998). In fact, Hoch and Loewenstein found that difficult items most often led to *greater*, not less hindsight bias. The confusion may have occurred because difficult items were also found to reduce participants' *overconfidence*—a related, but distinct measure of cognitive bias.

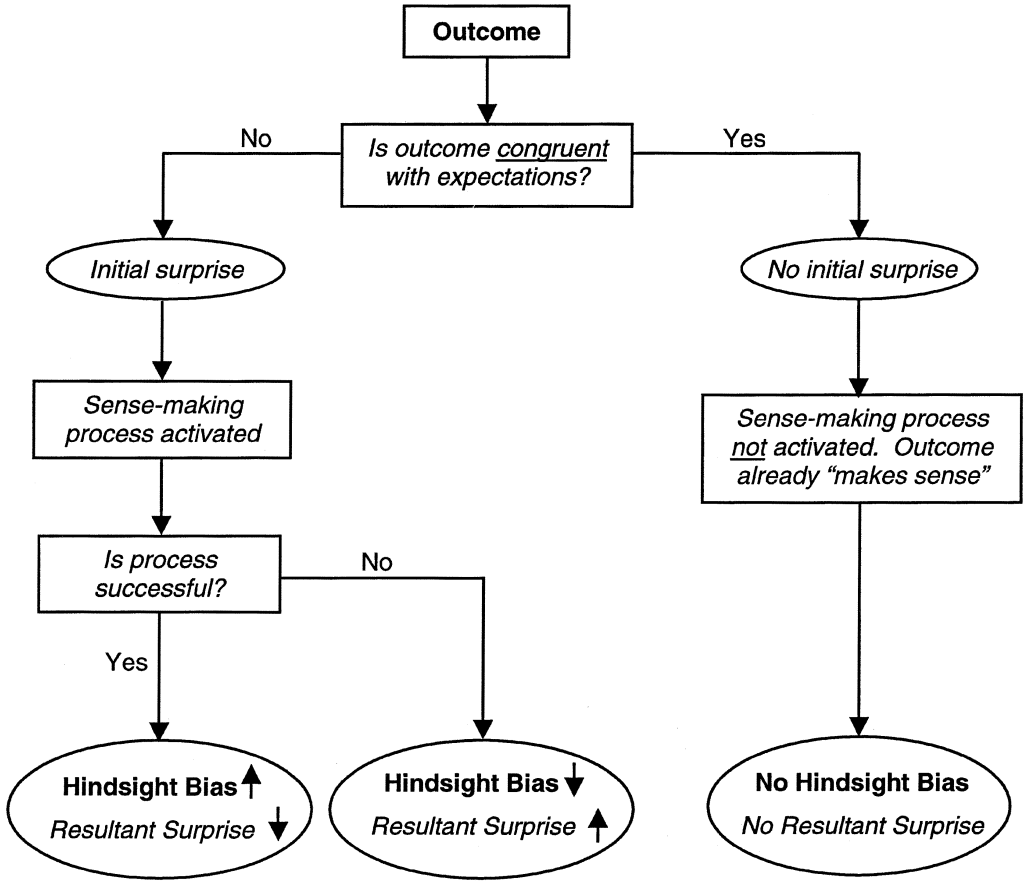

Figure 1. A sense-making model of hindsight bias. Hindsight bias is produced by making sense of unexpected outcomes. It is significantly reduced by unexpected outcomes that do not make sense, and removed by expected outcomes that do not require sensemaking. Resultant feelings of surprise are negatively correlated with hindsight bias because they are reduced by successful sensemaking. Note that relatively few outcomes in the past literature have been completely expected.

gruent outcome (i.e., ''initially'' surprising) could ultimately produce different levels of resultant surprise depending on the effectiveness of the sense-making process.[3]

The model of this process is presented in Figure 1. It is hypothesised that unexpected outcomes spontaneously engage a sense-making process, which, if successful, will produce hindsight bias. Although there may be multiple ways of ''making sense'' of an outcome, the model assumes that the magnitude of the bias is a function of the relative ease with which causal antecedents are uncovered (Roese & Olson, 1996). If an outcome is relatively easy to make sense of, it will produce a hindsight bias, and should seem relatively unsurprising.

Note that the outcome will not necessarily seem *completely* unsurprising. Hindsight bias merely requires a difference between pre- and post-outcome likelihood estimates, not a feeling that the person ''knew it all along''. For example, if an initial likelihood judgement is only 5% for some event, but later is recalled as 15%, hindsight bias would be said to exist, but no one would claim to have ''known all along'' that the outcome was going to occur. Thus the outcome would still be perceived as surprising but, importantly, less so.

However, if an outcome produces an effortful search that is not successful, this should reduce, remove, or even reverse hindsight bias, and produce resultant surprise levels that are relatively high. Essentially, surprise should result from the growing awareness that the sense-making process is not reaching closure. Finally, the model also states that outcomes that are expected, or obvious,

[3] Ofir and Mazursky (1997) clearly state that ''acknowledged surprise'' is associated with the reversal of the hindsight bias (p. 52). What is less clear is whether they view this form of surprise as the cause, result, or independent of sense-making activity. The present model suggests that acknowledged surprise is likely the *result* of a (failed) sense-making process.

should produce modest (if any) hindsight and little surprise. Because the outcomes already make sense, the processing thought to produce hindsight bias is less likely to be activated (Schkade & Kilbourne, 1991).

The sense-making model can be applied to a number of findings in the hindsight literature. For example, Wasserman et al. (1991) found that providing ''chance'' explanations for an outcome (e.g., tornado) produced a smaller hindsight bias than providing typical ''deterministic'' explanations. Similarly, Tan and Lipe (1997) showed that outcomes that were perceived as uncontrollable (due to their unpredictability) produced smaller hindsight effects. Both results could be explained in the present model by noting that, because chance factors have no causal antecedents, they are probably difficult if not impossible to make sense of. Pohl (1998) also reported that certain types of feedback believed to be ''implausible'' might be less susceptible to hindsight effects.

Roese and Maniar (1997) provide evidence for two aspects of the model. First, they found that in games for which Northwestern's football team was expected to lose, their (unexpected) wins produced large hindsight effects in fans. However, by the third game of their study, as expectations for success had grown considerably, the team's win produced no hindsight effect at all—presumably because the now-expected win required little sense-making activity. In some conditions, however, they induced participants to engage in causal thinking by either (a) having them ''go over the game over and over in their heads'', or (b) by thinking ''if only'' and imagining how the game might have turned out differently. Both techniques produced an increase in causal thoughts, and both produced significant increases in the hindsight bias for all games (see also Roese & Olson, 1996, Study 1).[4]

An examination of Ofir and Mazursky's (1997) stimulus materials suggests that in all three of their experiments they presented outcomes that appear to be very difficult to make sense of. In Experiment 1, for example, a medical patient is said to die, but, similar to Wasserman et al.'s (1991) study, there are no causal antecedents (other than chance) that would explain why the patient would die. Although the very unexpected death is sure to engage intense sense-making activity, it is unlikely that this activity will be successful because all of the evidence points to the patient living. Further, this process could produce the ''reverse'' hindsight bias that Ofir and Mazursky found. Because the only information available to participants points to the alternative outcome (patient lives), the increased processing caused by the unexpected outcome will produce likelihood estimates for the expected outcome that are even higher than they were before it was found that the opposite outcome occurred.

Although Ofir and Mazursky (1997) propose a similar mechanism, the current model may present a more parsimonious approach. Ofir and Mazursky claim that very surprising outcomes activate additional causal processing that, when combined with acknowledged surprise, offsets or reverses the effects of typical hindsight processing. Rather than invoke two competing processes, the present model suggests that only one process (sense-making) is at work and that this process must be successful for hindsight bias to occur, regardless of how initially surprising the event may be.

It might be worth considering what role defensive processing would play in the proposed model. Typically, we respond to threatening outcomes by forming excuses to reduce our perceived responsibility for the outcome (Markman & Tetlock, 2000). Such behaviour has been shown to reduce hindsight for measures that imply responsibility—such as the ''foreseeability'' ratings used by Mark and his colleagues (Mark & Mellor, 1991, Mark, Boburka, Eyssell, Cohen, & Mellor, 2003-this issue). Consider this, however: according to Snyder and Higgins (1988) much of excuse making involves shifting attributions for negative personal outcomes from internal (''I'm incompetent'') to external (''I'm overworked and tired''). One intriguing possibility is that although shifting from internal to external causes may reduce perceived blame for an event, it does not necessarily reduce the perceived *likelihood* of the event. In fact, a search for external causes for an event should still amass more causal antecedents for that outcome than if one had not tried to make excuses. The ironic result may be that although defensive pro-

[4] Note that ''if only'' counterfactual processing is different than the ''consider the opposite'' approach (Arkes et al., 1988; Slovic & Fischhoff, 1977). The latter approach asks people to consider a *different* outcome occurring under the *same* conditions. This weakens the causal link between the actual outcome and pre-existing conditions. ''If only'' processing, however, asks people to consider a *different* outcome occurring under *different* conditions. This strengthens the causal link between the outcome and pre-existing conditions, producing greater hindsight. When spontaneously making sense of an outcome, this form of ''if only'' processing is typically invoked (Roese & Olson, 1996).

cessing might help to reduce one's sense of culpability, it might also increase hindsight bias as measured by likelihood ratings (cf. Roese & Maniar, 1997). Such a possibility presents interesting directions for future research.

THE PRESENT RESEARCH

The present research will test predictions concerning the *absence* of hindsight bias derived from currently competing theories and from the proposed sense-making model. The primary goal of this research is to untangle the seemingly disparate predictions concerning outcome ''surprisingness''. A secondary goal of this research is to examine the limits of the defensive-processing hypothesis. Although defensive-processing models appear to imply the stipulation of culpability (cf. Mark et al., 2003-this issue), this has not been directly tested. Does defensive processing occur for upsetting outcomes for which one is not responsible?

The first of the present experiments compares predictions of two groups of loyal college basketball fans—one in foresight and the other in hindsight—to test the effects of both outcome expectedness and self-relevance in a real-world setting. The second and third experiments manipulate self-relevant outcome information via a bogus ''cognitive abilities test'' to examine these same effects, but in a laboratory setting. A fourth experiment presents relatively unexpected psychological research outcomes that are not self-relevant, but are either relatively difficult or easy to understand.

EXPERIMENT 1

Sports contests have proven to be successful outlets for testing hindsight predictions (e.g., Leary, 1981; Roese & Maniar, 1997). In the present research, the availability of both home and visiting fans provides an opportunity to examine the effects of both naturally occurring differences in outcome expectations and the self-relevant implications of those outcomes. First, presumably home fans would think that a home team win was more likely than would visiting fans. Second, such a win has different connotations for home versus visiting fans.

Roese and Maniar (1997) found that participants induced to generate causal explanations for football game outcomes exhibited the greatest hindsight bias. Because unexpected sports outcomes have been shown to spontaneously produce greater sense-making activity (Lau & Russell, 1980), it is predicted that whichever team receives an unexpected outcome (regardless of whether it is a win or loss) will exhibit greater hindsight bias. Conversely, if either team receives an outcome that is expected, they should exhibit no hindsight bias. And if either team's fans receive a very unexpected outcome—particularly one that is difficult to make sense of—these fans are also not expected to exhibit the bias.

What predictions might a defensive-processing hypothesis make? Although sporting events do not produce outcomes for which fans feel any responsibility, Cialdini et al. (1976) have shown that sports can have a significant effect on fans' self-esteem even though fans play no role in determining the outcome. Leary (1981) predicted that greater self-esteem concerns (e.g., public responding, high ego-involvement) would *increase* the hindsight bias for college football fans, because it is self-serving to act as though one ''knew it all along'' what would happen. However, his results showed no effect of such motives on likelihood estimates. Of course, defensive processing might still have played a role, however, because this study did not record whether participants were home or visiting team fans, and thus could not take into account the self-relevant ''meaning'' of the outcome. Roese and Maniar (1997) also did not distinguish between home and visiting fans, noting that ''nearly all'' of their participants were home team fans (p. 1251). Had the distinction been made, the defensive-processing hypothesis might predict that fans of the losing team would exhibit less hindsight bias than fans of the winning team. If such an effect were not found, this could suggest that culpability is a necessary condition for activation of the defensive-processing mechanism.

Method

We approached 281 people immediately before or after three college basketball games. The home team (Wake Forest) played three different visiting teams (Duke, University of North Carolina, and North Carolina State). College basketball was chosen because of strong fan involvement. After indicating their team allegiance, participants were asked ''Using a value from 0–100%, how likely is it (was it) that the home team, Wake Forest, will

(would) win this game?''. Hindsight bias was defined as the difference between pre- and post-game estimates. A between-subjects design was used in that different participants were (randomly) assigned to the pre-game and post-game conditions. During this season, Wake Forest had 23 wins and 5 losses, and won all three of the games in the study. Assistants helping to collect data observed that the losing team fans in all three games appeared visibly upset by their team's failure.

Results

A three-way ANOVA was performed on participants' likelihood judgements with team (Duke, NC State, or UNC), fan (home or visiting), and time (response given before or after game) as the independent variables. Table 1 shows the mean, median, and standard deviation values for the three teams. No predictions were made concerning the team variable, and it did not significantly interact with any other variables, all ps > .51. Thus, in order to increase power, all further analyses collapsed across this variable. Visiting fans were significantly less likely (M = 44%) than home team fans (M = 82%) to believe in foresight that Wake Forest would win the game, $F(1, 269)$ = 237.8, p = .001, partial η^2 = .47. Most important was a significant two-way interaction of fan × time, indicating that the home and visiting fans exhibited hindsight biases of different magnitudes, $F(1, 269)$ = 4.78, p = .03, partial η^2 = .02. Planned comparisons showed that although the visiting team fans exhibited marginally significant hindsight effects, $F(1, 269)$ = 3.51, p = .07, the home team fans did not exhibit any hindsight bias, p = .43.[5]

TABLE 1
Estimated likelihood of home team win (0–100%)

		Game		
Condition		*Duke*	*NC State*	*UNC*
Home fans				
Foresight	*Md*	80	85	75
	M	82.7	85.1	78.6
	SD	11.1	9.5	17.8
	n	21	30	27
Hindsight	*Md*	80	85	77
	M	80.9	82.4	74.9
	SD	12.2	12.2	23.7
	n	21	28	22
Opposing fans				
Foresight	*Md*	40	50	38
	M	38.8	47.0	32.5
	SD	15.8	22.9	26.8
	n	17	32	26
Hindsight	*Md*	50	65	50
	M	49.3	52.7	40.0
	SD	12.2	33.3	25.7
	n	15	19	23

Discussion

The present study indicates that, under certain conditions, salient real-world outcomes can fail to produce any significant hindsight bias. Because home and visiting team fans gave different foresight estimates (82% and 44%, respectively), this study provides an opportunity to examine the effect of varying levels of unexpectedness on the hindsight bias. In particular, home team college basketball fans, who expected their team to win each game, did not exhibit a hindsight bias. Visiting team fans, however, who did *not* expect Wake Forest to win, exhibited a marginally significant hindsight bias.[6] These results are consistent with the model's prediction that relatively unexpected outcomes produce hindsight bias, but expected outcomes do not (cf. Cannon & Quinsey, 1997, Study 1). However, the results do not appear to support Ofir and Mazursky's (1997) main prediction that surprising outcomes will produce reverse hindsight bias. Of course, this may be because the outcomes in this study were not sufficiently surprising to invoke any additional processing to offset the bias. Recall, however, that Ofir and Mazursky also predict that unsurprising outcomes *will* produce hindsight. This did not occur in the present study. The highly expected win for the home team fans failed to produce any hindsight bias.

One concern in this study is that the lack of hindsight bias for the home team fans could be caused by ceiling effects. Although it is true that the home team fans gave relatively high estimates in foresight (M = 82%, Md = 83%), there is arguably still ''room'' for fans asked in hindsight

[5] The same analyses using non-parametric tests produced virtually identical p-values: .08 and .55, respectively.

[6] As one reviewer notes, the NC State fans may have reduced this effect somewhat. These fans were not as confident that their team would win, and probably did not find their team's loss to be as unexpected—and thus exhibited a smaller hindsight bias.

to give even larger estimates. This did not occur, however, and in fact their mean and median likelihood estimates for a win indicated a small ''reverse'' hindsight bias for all three games, although not a statistically significant one. Thus, although a ceiling effect cannot be completely ruled out, it seems unlikely.

Finally, consider the results from a motivational perspective. This study indicates that a negative outcome, even for dedicated college basketball fans, is not sufficient to remove the hindsight bias. This may present a limitation for the defensive-processing hypothesis. That is, a sense of culpability for a negative outcome, rather than just self-relevance, may be necessary for the defensive-processing mechanism to be activated. If feelings of culpability do not exist, then the process at work may be primarily cognitive in nature (Fischhoff, 1977; Schkade & Kilbourne, 1991).

EXPERIMENT 2

Use of real-world outcomes in the first experiment produced some design limitations. First, although the college basketball games used in Experiment 1 did produce strong emotional reactions in fans and could even have had implications for those fans' self-esteem, no one could claim that the fans were responsible for the outcomes. Second, participants were not randomly assigned to each outcome condition. Third, no explicit measures of sensemaking, defensive thinking, or resultant surprise were included. Experiment 2 attempts to resolve these issues, in a laboratory setting, by randomly assigning participants to receive either positive or negative test-performance feedback, for which they would presumably feel responsible, and recording both participants' thoughts about the outcome and their resultant surprise at it.

Participants in Experiment 2 received bogus performance feedback designed to be either threatening or not threatening to participants' self-image. Outcome expectations were manipulated by giving feedback that was either consistent or inconsistent with participants' a priori belief about their own performance abilities. An explicit measure of resultant surprise was included, and an attempt to measure sense-making activity was achieved by coding participants' thoughts as either consistent or inconsistent with the outcome. Defensive processing was measured by counting the number of excuses participants provided. As an additional measure of the effects of defensive processing, participants' responses were compared to those of an additional set of (yoked) control participants, who had no personal investment in the outcomes, and thus would not be expected to exhibit such processing.

Method

Participants. Introductory psychology students were screened for academic self-esteem (Fleming & Courtney, 1984). Students scoring in either the upper or lower third of the distribution were invited to participate in the experiment on the condition that they bring an acquaintance who was not enrolled in introductory psychology. A total of 145 student–acquaintance pairs participated.

Procedure. All participants were given a ''Cognitive Abilities'' test, and led to believe that it was an excellent predictor of college performance. Much of the bogus test was modelled after items in Raven's Progressive Matrices (Raven, Raven, & Court, 1995), although some additional word problems were included. After initial instructions were given to all participants, the acquaintance was taken to a separate room ostensibly to create a less distracting test environment for both participants. In fact, this room had a one-way mirror that allowed the acquaintance to secretly observe the actor as he or she took the test. The actor was positioned close enough to the mirror that the observer could read the test and see the responses that the actor gave. In addition, the observer was provided with his/her own copy of the test. Thus, the observer served as a yoked control—able to read along and see the answers given by the actor, but with no personal investment in the outcome. Similar results for both actors and observers in a given condition (deemed threatening to actors) would weaken any claim of defensive processing. Actors were given 20 minutes to complete the test. The experimenter then removed the test and answer sheet, ostensibly to score it on a computer.

Outcome feedback. Five minutes later the experimenter returned with a printout indicating the participant's percentile rank on the test compared to other students at the university. Actors were either given negative feedback (30th percentile), positive feedback (90th percentile), or no feedback (control). The observer, who remained in the adjacent room during this time, was also shown a copy of the actor's feedback sheet.

Thought listing. Participants were then given five minutes to record any thoughts (one per line) about their test performance and feedback. Observers listed thoughts about the actor's performance. Participants were then asked to indicate next to each thought whether it was suggestive of a good performance, a poor performance, or irrelevant.[7] Irrelevant thoughts were not analysed. Some examples of reasons for a poor performance in this study were:

I'm horrible with word problems!
Other people are able to see patterns better than me.
My SAT scores are pretty consistent with this (bad outcome).

Examples of reasons for a good performance in this study were:

I've always been good at this stuff.
The problems seemed easy.
I checked my answers twice—I think they're mostly right.

Hindsight measure. Actors given no performance feedback were asked to indicate, ''What percentile rank did you think you would fall into right after you completed this test?'' Actors given feedback were asked to indicate how they thought they would have responded *before* receiving feedback. Observers also gave estimates for the actor they observed.

Resultant surprise. Actors and observers receiving outcome information were asked to rate their resultant surprise at the outcome using a 9-point scale with anchors of *not at all* and *extremely* surprising. The presentation order for surprise and hindsight measures was counterbalanced to control for carryover effects.

Additional measures. Actors and observers were asked, also using a 9-point scale, to rate the following items: (a) test difficulty, (b) their general test-taking abilities, (c) the other person's general test-taking abilities. These questions appeared randomly in between surprise and hindsight ratings.

Suspicion. After completing all of the measures, all participants were asked to indicate whether or not they were suspicious of any of the procedures in the experiment, particularly the test and feedback they received. If either the actor or the observer in a yoked pair indicated moderate or greater suspicion, the pair was not included in the analyses described below.

[7] The ''self-coding'' approach was used because participants had more insight into the meaning of their idiosyncratically written thoughts than we did. Later, an independent coder, blind to condition, recoded the thoughts. Agreement with participants' own categories was 93%.

Results

A total of 19 actor/observer pairs were dropped from the analyses. Approximately half of these were due to record-keeping errors or failure to complete the questionnaire. The other half were due to suspicion of the procedures, usually voiced by observers. There remained 126 pairs of participants.

Manipulation checks. Regardless of feedback, low self-esteem actors rated their test-taking abilities lower than did high self-esteem actors (Ms = 4.9 vs 6.5, respectively), $t(124) = -5.6$, $p < .001$, $d = 1.0$. Actors receiving negative feedback rated the test more difficult ($M = 6.5$) than actors receiving positive feedback ($M = 5.5$), $t(80) = 2.62$, $p = .011$, $d = 0.57$. Control actors with low self-esteem expected to perform worse than did those with high self-esteem, $t(42) = 3.05$, $p = .004$, $d = 0.94$. We may surmise, then, that low self-esteem actors found the positive outcome to be more unexpected and that high self-esteem actors found the negative outcome more unexpected. Finally, all observers rated the test to be moderately difficult ($M = 6.2$) and rated the actors to have moderate test-taking abilities (M = 6.5), regardless of which condition they were in.

Basic hindsight measure. Table 2 presents cell means, medians, and standard deviations for actors' and observers' estimates of the actors' performance on the cognitive abilities test (in percentile units).

Actors. Four planned-comparisons were performed comparing both positive and negative outcome conditions with their same-self-esteem controls. The hindsight bias was not found for outcomes that were generally congruent with participants' expectations, but was found for those that were relatively incongruent. Predictions made by low self-esteem actors receiving a negative outcome did not differ from those made by participants in the control condition, $p = .89$, but participants receiving the positive outcome gave significantly larger estimates than controls, $p =$

TABLE 2
Estimates of actors' test score (percentile rank) as a function of actor self-esteem (SE) and outcome

		Outcome		
		Negative (30%)	*Control*	*Positive (90%)*
Actors' estimates				
Low SE	*Md*	50	50	70
	M	49.1[a]	49.7[a]	65.1[b]
	SD	13.2	15.4	17.4
	n	22	22	21
High SE	*Md*	50	60	70
	M	51.7[a]	64.1[b]	68.3[b]
	SD	18.1	16.0	16.4
	n	19	22	20
Observers' estimates				
	Md	50	70	80
	M	51.0[a]	65.3[b]	74.9[c]
	SD	11.3	14.9	13.4
	n	41	44	41

Estimates range from 0 to 100%. All comparisons are made within a self-esteem condition (i.e., row) across outcomes. Cells sharing a superscript do not differ significantly at the .05 level.

.002, $d = 1.22$. For high self-esteem actors, this pattern was reversed. Those receiving the positive outcome did not differ from the control group, $p = .42$, but those receiving the negative outcome gave significantly lower estimates, $p = .014$, $d = 0.73$.

Observers. The outcomes used in this experiment were not self-relevant for observers. Although some actor/observer pairs were friends, the majority were mere acquaintances. Thus, observers were not expected to be aware of actors' self-esteem concerns, nor were they expected to have much personal stake in actors' test results. As predicted then, there was no main effect of self-esteem, nor any interaction with self-esteem and outcome, ps = .72 and .46, respectively. Values in Table 2 are collapsed across self-esteem for observers. The results indicate typical hindsight biases. Compared to the control condition, observers receiving negative feedback (about actor's performance) gave lower performance predictions, $p < .001$, $d = 1.1$, and those receiving positive feedback gave higher performance predictions, $p = .003$, $d = 0.68$.

Resultant surprise ratings. Actors. All actors reported moderate levels of surprise. As might be expected, low self-esteem actors reported being significantly more surprised by positive outcomes ($M = 7.1$, $SD = 1.5$) than by negative outcomes ($M = 5.7$, $SD = 2.4$), $t(41) = 2.28$, $p = .03$, $d = 0.69$. However, high self-esteem actors did not differ significantly in their reported surprise levels for positive ($M = 6.2$, $SD = 2.1$) and negative outcomes ($M = 5.2$, $SD = 2.4$), $p = .19$.

Observers. Observers reported that the negative outcome was more surprising ($M = 6.5$, SD = 1.8) than the positive outcome ($M = 4.8$, $SD = 2.74$), $t(80) = 9.53$, $p = .003$, $d = 0.70$. No other effects were significant, Fs < 1.

Correlations between hindsight magnitude and resultant surprise. The two primary predictions from the model are that (a) hindsight bias would be greatest in conditions with unexpected outcomes, and that (b) resultant surprise and hindsight bias would be inversely related. Analysis of the likelihood estimates for both actors and observers provides support for the first prediction. To determine correlations between surprise and hindsight bias, however, a single value indicating the magnitude of the hindsight bias must be calculated for each participant. This was achieved by subtracting the prediction of each participant receiving an outcome from the mean of his or her (high or low self-esteem) control group. The resulting value was multiplied by -1 for positive outcomes so it would be positive for typical hindsight effects, but negative for a reverse hindsight effect. The same procedure was used with observers, although their control condition did not separate into high and low self-esteem.

As predicted, increases in hindsight were associated with decreases in resultant surprise, $r_{\text{actors}} = -.56$, $r_{\text{observers}} = -.47$, both ps < .001. Similar correlations were found within each of the four cells created by the 2 × 2 crossing of self-esteem and outcome valence, all ps < .02. That is, regardless of whether outcomes were generally congruent or incongruent, people who found them to still be surprising after 5 minutes of thought showed less hindsight bias. This result is consistent with Ofir and Mazursky's (1997) finding that surprising outcomes produce little or reverse hindsight bias. Indeed, scatter plots of hindsight magnitude and resultant surprise ratings for each cell (not shown) consistently indicate that participants expressing the most surprise showed a "reverse" hindsight bias, in that their hindsight magnitude scores were negative.

Thought listing procedure. Thought listings were included because they might provide an indication of the ease with which participants

made sense of the outcomes. For example, thoughts consistent with the outcome (e.g., reasons for poor performance given after negative outcome) might indicate successful sense-making activity, while inconsistent ones might indicate sense-making difficulty.

Actors. Low and high self-esteem actors did not differ in the total number of thoughts listed (*M*s = 7.3 and 7.6, resp.), $p = .18$. Table 3 reports the *proportion* of total thoughts of each type (e.g., reasons for good or poor performance) in each cell. Positive outcomes produced no significant differences in thoughts compared to those reported by control participants. Negative outcomes, however, exerted a significant and similar effect on both low and high self-esteem participants. Compared to same-self-esteem controls, participants receiving negative outcomes thought of significantly more reasons why they would perform poorly, and significantly fewer reasons why they would perform well. Although such a finding is consistent with the model for high self-esteem participants—they showed hindsight bias for negative outcomes, but not for positive outcomes—it does not appear to explain the opposite results of low self-esteem participants.

Perhaps a better way to detect the effects of sensemaking is to examine within-cell correlations between hindsight magnitude, surprise, and the proportion of outcome-consistent and inconsistent thoughts. We would expect the proportion of outcome-consistent thoughts (e.g., reasons for good performance following positive feedback) to be positively correlated with hindsight magnitude and negatively correlated with resultant surprise. Conversely, we would expect exactly the opposite for inconsistent thoughts—they should be negatively correlated with hindsight magnitude and positively correlated with resultant surprise.[8] Indeed, this was found for high self-esteem actors in all conditions. All eight predicted correlations were in the expected direction with magnitudes of $|r| \geq .47$, all ps $< .03$. Again, however, the results were less clear for low self-esteem participants. Although all correlations were in the expected direction, only three of the eight approached significance. First, as predicted, hindsight magnitude was negatively correlated with inconsistent thoughts but only for positive outcomes, $r = -.53$, $p = .02$. Further, and as predicted, resultant surprise was negatively correlated with outcome-consistent thoughts, $r = -.57$, $p = .01$, and marginally so with outcome inconsistent thoughts, $r = .40$, $p = .07$, but each of these only occurred for negative outcomes.

Observers. Similar findings were obtained for the thought listings of observers. Except for the negative outcome, consistent thoughts were positively correlated with hindsight bias and negatively correlated with resultant surprise, while inconsistent thoughts showed the opposite effect, all rs = .33 or greater, all ps $< .035$. In the negative outcome condition, hindsight magnitude was not significantly correlated with either type of thought, both ps $> .30$.

Thus, for both actors and observers, significant correlations between thoughts and hindsight (where they were found) indicated that people who spontaneously generated more outcome-inconsistent thoughts exhibited reduced hindsight bias—a finding consistent with past literature (e.g., Arkes et al., 1988), and with the present model. It is unclear, however, why these correlations were not consistently obtained, although one possible explanation will be discussed in the following section.

TABLE 3
Proportion of thoughts predictive of good and poor test score as a function of self-esteem and outcome

		Outcome		
		Negative	*Control*	*Positive*
Low self-esteem				
Good score	*M*	.29[a]	.45[b]	.43[b]
	SD	.19	.17	.23
Poor score	*M*	.67[a]	.54[b]	.46[b]
	SD	.18	.17	.22
High self-esteem				
Good score	*M*	.32[a]	.53[b]	.54[b]
	SD	.17	.16	.16
Poor score	*M*	.59[a]	.46[b]	.45[b]
	SD	.21	.16	.14

Values in the same row with different superscripts are significantly different, $p < .05$. Column values for a given self-esteem condition do not add to 100% because irrelevant thoughts are excluded.

[8] To do this, thought listings were simply recoded as either consistent or inconsistent with the actual outcome, and control subjects were dropped from the analysis. Note that 8 separate within-cell correlations arise from a 2 (thought consistency) × 2 (outcome) × 2 (hindsight or surprise measure) crossing. Multiplying this by 2 self-esteem conditions yields a total of 16 within-cell correlations. Space prevents US reporting these in a separate table.

Discussion

Experiment 2 was designed to test predictions concerning both defensive processing and the unexpected nature of outcome information on the presence of the hindsight bias.

Defensive processing. Defensive processing is believed to be activated by people's desire to reduce their sense of culpability for a negative outcome (Louie et al., 2000; Mark & Mellor, 1991). Thus, the defensive-processing mechanism may only be activated when the participants could feel responsible for causing the outcome they receive. Although participants in Experiment 1 would not be expected to experience such feelings, those in Experiment 2 should feel responsible for their own good or poor performance on a ''cognitive abilities'' test. To the extent that this is true, the defensive-processing hypothesis predicts that participants in both low and high self-esteem conditions would deny the foreseeability of any negative feedback they get about their test performance—these people should exhibit little or no hindsight bias for negative feedback. It is also reasonable to predict that positive self-relevant outcomes will produce typical hindsight bias because it is to participants' advantage to enhance the foreseeability of their causing such events (Louie et al., 2000). Finally, yoked control participants should exhibit typical hindsight bias for both negative and positive outcomes, because neither outcome has implications for their self-image.

The data do not provide unequivocal evidence for a defensive-processing mechanism. Although low self-esteem participants did not show the bias when given negative performance feedback (as predicted by defensive processing), high self-esteem participants *did* exhibit a significant hindsight bias for negative feedback. These data might be taken to indicate an interesting qualification of the defensive-processing hypothesis: Perhaps only low self-esteem individuals utilise defensive processing when faced with threatening outcomes. Although previous research indicates that it is usually high, rather than low, self-esteem individuals who are most likely to employ defensive mechanisms in the face of threatening outcome information (Agostinelli, Sherman, Presson, & Chassin, 1992; Brown & Gallagher, 1992), this possibility should be examined.

To test this idea, two independent raters counted the number of reasons given by actors for their poor performance that could be classified as ''excuses''—thoughts that shifted blame for a poor performance from internal to external causes (e.g., ''This desk is too small to write on'', ''That fan is distracting'', etc.) (cf. Markman & Tetlock, 2000). Inter-rater agreement was high (89%) and discussion resolved any differences. High and low self-esteem participants did not differ in the proportion of total thoughts classified as excuses, (*M*s = 50% and 52%, respectively), $p = .78$. Further, there were no significant correlations between number (or proportion) of excuses and either hindsight magnitude or resultant surprise ratings for either high or low self-esteem participants, all ps $> .31$. Thus, although it is still possible that low self-esteem participants could have reduced the hindsight bias for negative outcomes via defensive processing, this study was unable to uncover any indication of such processing. It may also be that defensive processing is most likely to occur with questions that imply culpability. For example, most (though not all) studies that have shown reductions in hindsight effects for threatening outcomes asked participants to indicate how ''foreseeable'' the outcome was rather than provide numerical estimates of the likelihood of the outcome (cf. Mark et al., 2003-this issue; Renner, 2003-this issue).

Another concern for the defensive-processing hypothesis, however, is that positive outcomes did not produce hindsight bias for high self-esteem individuals. Again, one might predict that any outcome reflecting positively on one's self-image would produce a hindsight bias (Louie et al., 2000). Thus, it appears that something more than mere self-relevance and outcome valence plays a role in determining whether or not the hindsight bias will occur.

Outcome expectations. Regardless of valence, outcomes that were generally congruent with a priori expectations produced no hindsight bias, whereas those that were incongruent produced sizeable hindsight biases. In particular, participants with low academic self-esteem showed the bias for positive outcomes, but not for negative outcomes, and those with high academic self-esteem showed the bias for negative outcomes, but not for positive outcomes. These results are consistent with a cognitive mechanism of hindsight bias in which incongruent events—independent of their implications for the self—produce greater sense-making activity, which, in turn, produces greater hindsight bias (cf. Schkade &

Kilbourne, 1991). The data are also consistent with another prediction from the model, namely that hindsight bias and resultant surprise will be inversely related. Those who found the outcomes to be most surprising after spending 5 minutes thinking about them, showed the least, and often reverse hindsight bias (cf. Ofir & Mazursky, 1997).

That the hindsight bias in the low self-esteem condition (positive outcome) was considerably larger than in the high self-esteem condition (negative outcome) may simply reflect a difference in the degree to which the outcome feedback was incongruent with expectations for the two groups. When compared with estimates from their own control condition it appears that the 90th percentile (positive) feedback for low self-esteem participants was somewhat more unexpected (90% − 50% = 40%) than the 30th percentile (negative) feedback was for the high self-esteem participants (65% − 30% = 25%). Data from Experiment 3 will help to verify this assertion. Similarly, the differences in hindsight magnitudes for observers is probably due to the negative outcome feedback being somewhat more different from control group expectations than the positive feedback.

Additionally, the fact that the hindsight bias for positive outcomes was significant for observers but not for high self-esteem actors (who gave similar foresight estimates) might seem difficult to explain. However, one might counter this concern by recalling that the magnitude of the bias for observers was greater for negative outcomes than for positive, and that this is consistent with results for high self-esteem actors. Another possibility is that modesty on the part of high self-esteem actors artificially reduced their foresight estimates. If this were the case, then high self-esteem actors would not actually be as (initially) surprised as the observers by their positive feedback, and the hindsight bias of the observers, but not actors, would make perfect sense. Of course, this explanation is speculative, and these results clearly invite further study. Despite the lack of a definite cognitive explanation, however, note that defensive processing also cannot explain this apparent discrepancy, unless the positive outcome could be viewed in some way as threatening to the high self-esteem actors.

Although the data in this experiment are consistent with the predictions of the sense-making model, the experiment does not manipulate or measure sense-making activity directly. Thought listings were included because it was believed that they might provide an indication of the ease with which participants made sense of the outcomes. A greater proportion of outcome-consistent thoughts might indicate successful sense-making activity, while a greater proportion of inconsistent thoughts might indicate difficulty in sensemaking. However, Schwarz (1998) has suggested that a distinction should be made between what thoughts (or how many) come to mind and how easily those thoughts come to mind. He showed that it is often the *ease* with which thoughts come to mind, rather than the absolute amount of thoughts, that has the biggest impact on judgements. Unfortunately, the measures used in the present study do not provide any indication of the relative ease with which various thoughts were brought to mind—only how many came to mind (see also Sanna, Schwarz, & Stocker, 2002). This may explain the sometimes inconsistent correlations between thoughts and hindsight and surprise measures, and should be examined in future research.

EXPERIMENT 3

Experiment 2 measured *resultant* surprise and showed that it is inversely correlated with hindsight bias. But it is also argued that differences in hindsight (and resultant surprise) are produced primarily by differences in outcome expectedness or *initial* surprise. Although in Experiment 2 expectation levels were inferred from participants' academic self-esteem (Fleming & Courtney, 1984) and from the control condition likelihood estimates, it seems instructive to also obtain a more direct measure of initial surprise. Thus, Experiment 3 asked participants to provide surprise ratings *immediately* after receiving the outcome, rather than after 5 minutes of thought.[9] It was predicted that given little time to make sense of the outcome, those in the two incongruent outcome conditions would provide greater ratings of surprise than would those in the two congruent outcome conditions. Thus, this experiment serves as a manipulation check for outcome incongruence in Experiment 2.

[9] Because people are likely to process outcome information the moment they receive it, it may be impossible to perfectly capture "initial" surprise. However, asking immediately after participants receive the outcome is arguably more accurate than asking 5 minutes later.

Method

A total of 44 introductory psychology students of high and low academic self-esteem were run in small groups of between one and four, with no observers. They did not speak to each other and were not aware of the feedback other participants received. Administration of the cognitive abilities test was identical to Experiment 2. After receiving feedback, participants were immediately given a short questionnaire that first asked how surprised they were by the results. They were then asked additional filler questions, asking them to indicate why they felt as they did, and other questions that were used in Experiment 2 (e.g., performance estimate, test difficulty, etc.). A no-outcome condition was not included, as the goal was to measure initial feelings of surprise at the outcome.

Results and discussion

The goal of this experiment is to show that expectations for the outcome (i.e., ''initial'' surprise) differed across conditions in Experiment 2. Because the order of the surprise and other questions was not counterbalanced, only the results for surprise ratings are described below. The lack of a control condition prevented hindsight bias from being calculated, and perhaps more important, asking participants to consciously consider their surprise first was likely to have carryover effects.

There was a significant outcome × self-esteem interaction on ''initial'' surprise ratings $F(1, 40) = 23.07$, $p < .001$, $d = 1.50$. Simple effects tests showed that low self-esteem participants found the positive outcome more surprising ($M = 7.0$, $SD = 1.3$) than the negative outcome ($M = 3.3$, $SD = 2.0$), $F(1, 43) = 20.3$, $p < .001$, $d = 1.77$, and that high self-esteem participants found the negative outcome more surprising ($M = 7.0$, $SD = 1.7$) than the positive outcome ($M = 5.6$, $SD = 1.8$), $F(1, 43) = 4.42$, $p = .042$, $d = 0.99$. As expected, participants who received outcomes that were inconsistent with their pre-existing levels of academic self-esteem rated the outcome as more surprising than did those who received outcomes that were consistent with their academic self-esteem. Although participants could still have taken some time to ''make sense'' of the outcome before responding to the surprise question, this measure is likely to correspond more closely with outcome unexpectedness than the measure taken in Experiment 2. Thus, it can now be said with more conviction that the outcomes that did *not* produce hindsight bias in Experiment 2 were those that were initially the least surprising (cf. Schkade & Kilbourne, 1991).

An additional test showed that for the two a priori ''expected'' conditions, high self-esteem participants receiving positive outcomes expressed significantly greater initial surprise ($M = 5.6$) than did low self-esteem participants who received a negative outcome ($M = 3.3$), $t(21) = 2.80$, $p = .011$, $d = 1.2$. This is consistent with the finding from Experiment 2 that high self-esteem participants showed a trend towards hindsight bias for positive outcomes that was, though not significant, similar to that of observers, whereas low self-esteem participants showed no bias whatsoever.

One additional finding deserves attention. For some conditions, surprise levels seemed to increase somewhat during the 5 minutes. For example, low self-esteem participants who received negative outcome gave an ''initial'' surprise rating of $M = 3.3$ in Experiment 3, but other low self-esteem participants, given time to think about this same outcome, actually rated it as somewhat more surprising 5 minutes later, $M = 5.73$ (Experiment 2). A similar shift occurred for high self-esteem participants receiving positive outcomes. Such an upward shift across time in the same condition is counterintuitive; surprise levels should generally decrease as one makes sense of the outcome. One explanation is simply that, because different people participated in the two experiments, some discrepancies in the absolute values are to be expected. A more reasonable possibility comes from the fact that many participants in Experiment 2 were run in the very beginning of their first year in school, whereas those in Experiment 3 were run later in that school year. It could be that academic experiences during the school year caused an amplification of participants' expectations—those with low academic self-esteem got lower, and those with high self-esteem got higher. Thus, we would not expect either group to be quite as surprised by their results. Admittedly, this is speculative, but it may provide a possible explanation. In any case, the results of Experiment 3, taken alone, seem to indicate that outcomes defined as expected and unexpected were generally perceived so.

Experiments 1, 2, and 3 included variables to test for predictions made by the defensive-processing hypothesis. Although this hypothesis

makes important claims about when the hindsight bias does not occur, defensive-processing motives somewhat confound the effects of sense-making ''ease'' on the hindsight bias. Thus, the final experiment does not use self-relevant outcomes, nor does it ask people to list their thoughts in such a way that they may be motivated to engage defensive processing. Instead it directly manipulates sense-making ease by presenting scenarios with relatively unexpected outcomes that have been pilot tested to be either easy or difficult to make sense of.

EXPERIMENT 4

The sense-making model suggests that hindsight bias is determined by the relative ease with which sense can be made of unexpected outcomes. Sensemaking is essentially a search for causal antecedents, something that occurs spontaneously for outcomes that are incongruent with expectations (Hastie, 1984; Sanna & Turley, 1996; Weiner, 1985). The final experiment provides a test of two predictions of the sense-making model of hindsight bias concerning the ease in sense-making of relatively incongruent outcomes. First, outcomes that are relatively easy to make sense of will produce a hindsight bias, and will ultimately seem relatively unsurprising (Prediction 1). Second, outcomes that are difficult to make sense of will produce either no hindsight bias or reverse hindsight bias, and will remain relatively surprising (Prediction 2).

In this study, an incongruent outcome is created by describing psychological research with a fairly intuitive solution, but then providing a counterintuitive outcome as the ostensibly true finding (see Choi & Nisbett, 2000; Davies, 1987; Slovic & Fischhoff, 1977, for other studies using a similar technique). Thus, sensemaking is expected to be activated by all outcomes used in this study. However, half of the outcomes were designed and pilot tested to be relatively difficult to make sense of, whereas the other half were designed to be relatively easy to make sense of.

Method

Participants and procedure. A total of 92 students at Wake Forest University received course credit for their participation. Two students were dropped from the analyses for failing to provide complete data. All participants received a booklet with six (single-paragraph) scenarios, each describing a psychology experiment. Five of the scenarios listed three possible outcomes for the experiment, and one (divorce study) listed two possible outcomes. Participants in the ''foresight'' condition (n = 47) were asked to indicate the likelihood of each outcome using a percentage from 0–100% with the restriction that the estimates must sum to 100%. Participants in the ''hindsight'' condition (n = 43) were given the same instructions, but were also told which of the outcomes was the actual finding, and that the alternatives were known to be incorrect. Hindsight participants were asked to ignore the outcome information. Participants receiving an outcome were also asked to indicate how surprised they were by the findings, using a 1–7 scale. As in Experiment 2, the order of presentation of resultant surprise and likelihood measures was counterbalanced to control for carryover effects. Although the unexpected nature of the outcomes should cause sensemaking to occur spontaneously, participants were also asked to proceed slowly, and prompted to make sense of the outcomes by encouraging them to ''make sure you understand everything about the scenario'' before turning to the next page and answering any questions.

Materials. A list of the six scenarios can be found in Tables 4 and 5. These scenarios were written and pilot tested to have outcomes that were unexpected, but that also varied in the difficulty with which participants could ultimately make sense of them. Pilot participants were asked to give their reactions to the ostensibly true outcome for each scenario. The scenarios written to be difficult to make sense of were typically thought to be so by most (although not all) participants, and those written to be easy to make sense of were found to be so by almost all participants. For example, one of the *difficult* scenarios stated, ''of girls who actually have sex, those who feel most guilty about it are most likely to get pregnant.'' Most participants had difficulty providing any reasons why this would be true. An *easy* scenario stated that when deciding on who to date, ''for both men and women, the only factor that mattered was how physically attractive the person was.'' Although students acknowledged that they would typically have guessed that only men care about ''looks'', it was relatively easy for them make sense of the finding that looks matter to both men and

TABLE 4
Likelihood estimates and resultant surprise ratings for research outcomes that are *difficult* to make sense of

Research outcome	*Foresight*		*Hindsight*	*Surprise*		*Surprise* × *Hindsight correlation*
Teens who feel guilty are more likely to get pregnant	*M*	12.9	19.9	*M*	5.4	$r = -.77$
	SD	20.3	24.3	*SD*	1.6	$p < .001$
	Md	5	10			
American students are less overconfident than Chinese	*M*	18.7	18.6	*M*	5.3	$r = -.26$
	SD	13.1	12.9	*SD*	1.0	$p = .09$
	Md	20	20			
Democrats favour death penalty more than Republicans	*M*	32.0	30.5	*M*	4.9	$r = -.44$
	SD	18.5	17.0	*SD*	1.2	$p = .047$
	Md	30	25			

No comparisons between foresight and hindsight estimates were significant, all $ps \geq .20$.

TABLE 5
Likelihood estimates and resultant surprise ratings for research outcomes that are *easy* to make sense of

Research outcome	*Foresight*		*Hindsight*	*Surprise*		*Surprise* × *Hindsight correlation*
''Looks'' are equally important to both men and women	*M*	27.3	45.3	*M*	3.95	$r = -.64$
	SD	18.5	19.9	*SD*	1.6	$p < .001$
	Md	20	40			
Seminary students only help if they aren't late for a talk	*M*	40.0	55.5	*M*	3.43	$r = -.51$
	SD	26.5	27.2	*SD*	1.6	$p = .001$
	Md	40	60			
Couples who argue are more likely to get divorced	*M*	55.0	74.0	*M*	2.41	$r = -.54$
	SD	20.7	21.5	*SD*	1.4	$p = < .001$
	Md	60	80			

All foresight/hindsight comparisons were significant, $ps < .01$. Effect sizes ranged from $d = 0.60$ to 0.93.

women. Thus, the scenarios used should be sufficient to test the model.[10]

After the actual experiment was conducted, numerous participants volunteered to the experimenter that they had already learned about one of the six research findings in their psychology class. After confirming that this particular topic (and none of the others) had been covered in many of the introductory psychology classes contributing to the subject pool, this scenario was dropped from analyses and a new scenario (Democrats favour the death penalty) was run with additional participants, after additional pilot testing. A filler scenario, which was not analysed, was always presented first to capture the ''within-subject'' nature of the design used with the first set of scenarios. Due to time constraints only 42 additional participants (21 in foresight, 21 in hindsight) could be run in this condition. Below is a full example of a scenario written to be difficult to make sense of (political attitudes) along with three possible outcomes given to participants. The ostensibly true outcome was indicated by two asterisks.

> Researchers in psychology have long been interested in the relationship between political affiliation and opinions about various issues. It is important to be able to predict people's opinions without actually asking them. One issue in particular is that of capital punishment. In the United States, capital punishment is a hot topic, and a series of studies has uncovered the following

[10] Curiously, although many participants had difficulty generating any explanations for the ''difficult'' outcomes, they were also unable to generate more than one or two reasons for the ''easy'' outcomes, despite their claims that these outcomes made a lot of sense. This is consistent the idea that the ease with which reasons are generated is not necessarily indicated by the absolute number of reasons generated (Schwarz, 1998).

about Republicans and Democrats. In most states, it appears that:

Republicans are more likely than Democrats to favour capital punishment.

** Democrats are more likely than Republicans to favour capital punishment.

There is no difference between the two parties regarding capital punishment.

Results and discussion

Planned comparisons were used to test for hindsight bias for each of the easy and difficult scenarios. Consistent with Predictions 1 and 2, the three outcomes judged by pilot participants to be easy to make sense of produced a sizeable hindsight bias, whereas outcomes that were judged difficult to understand produced no hindsight bias (see Tables 4 & 5).[11] Although all of the outcomes in this experiment were at least somewhat unexpected,[12] the foresight estimates for the difficult condition were lower, on average, than for the easy condition—although not exclusively so.

Unexpected outcomes typically produce sense-making activity (Olson, Roese, & Zanna, 1996; Weiner, 1985), and this activity has been shown to produce greater hindsight bias (Roese & Maniar, 1997; Roese & Olson, 1996). Thus, we might predict that the outcomes in the difficult condition would produce even more sense-making activity—and thus hindsight bias—than those in the easy condition. In fact, these outcomes produced no significant hindsight biases, whereas the outcomes in the easy condition produced significant hindsight biases. This finding seems to indicate that the ease with which sense can be made of an outcome, rather than ''surprisingness'' *per se*, is the more important determinant of hindsight bias.

Indeed, even though participants could proceed at their own pace, and thus likely took the time to try understand the outcomes, those receiving difficult outcomes still remained quite surprised. This is consistent with the idea that they had difficulty in making sense of these outcomes. As in Experiment 2, ratings of surprise in this experiment indicate ''resultant'' surprise rather than initial surprise, although little change between the two types would be expected for difficult outcomes. Also, as in Experiment 2, resultant surprise measures were negatively correlated with likelihood ratings. Such a finding is consistent with the idea that those participants who found an outcome to be difficult to make sense of (regardless of whether it was designed a priori to be difficult) found the outcome to be more surprising, and exhibited less hindsight bias. Of course, this study did not measure sense-making activity (or its success or failure) directly. Future research should examine ways to obtain non-intrusive measures of such activity. Some possibilities are discussed below.

GENERAL DISCUSSION

The hindsight bias is considered a robust phenomenon that occurs in myriad domains. The current research, however, was designed to delineate those times when we do *not* experience hindsight bias. Using both real-world and laboratory settings, the present research examined claims arising from two distinct perspectives, one cognitive and one motivational, both of which predict specific absences of the hindsight bias.

The cognitive perspective: Outcome surprisingness

Interestingly, the results of this research provide support for both Ofir and Mazursky's (1997) ''surprise reduces the bias'' position and Schkade and Kilbourne's (1991) ''surprise increases the bias'' position, sometimes in the same study. Experiments 1 and 2, for example, both showed that the outcomes that were least expected produced the greatest hindsight bias, and those that were most expected produced little or no bias (''surprise increases hindsight bias''). In Experiments 2 and 4, however, significant *negative* correlations were found between surprise ratings and hindsight bias in all conditions—those outcomes that were most surprising produced the least hindsight bias (''surprise decreases hindsight bias''). The position taken here is that outcomes that are unexpected or what might be termed ''initially'' surprising invoke a sense-making mechanism that, if successful, produces hindsight bias and a reduced sense of ''resultant'' surprise. If the sense-making process is not successful, however, resultant surprise will remain relatively high and hindsight bias should lower. It is resultant

[11] Non-parametric tests performed on the six scenarios revealed identical results.

[12] Foresight ratings for the divorce scenario were fairly large because, due to experimenter error, this scenario provided participants with only two outcomes. The remaining five scenarios used three outcomes, which will generally lower the estimate for any given outcome.

surprise that seems most consistent with Ofir and Mazursky's ideas, particularly their claim that surprise must be acknowledged before hindsight bias can be reduced or reversed. However, the model does not require that we have conscious awareness of initial surprise. In fact, in many cases it seems likely that we are not aware, for otherwise it would be difficult for hindsight to ''creep'' up on us, as Fischhoff (1975) suggested. On the other hand, it is conceivable that people *could* be aware of their initial surprise and still show hindsight bias, albeit of a slightly different form. Consider the expression, from the quote in the beginning of the paper: ''I was surprised, but not that surprised. I mean, it makes sense.'' The position held here is that if an outcome can be understood in light of one's previous knowledge, then hindsight bias should occur. Rather than saying ''I knew it all along'', however, the acknowledged surprise might cause a person to say, ''I *should* have known it all along.'' To my knowledge no research has looked at the hindsight bias from this perspective.

For those who did not claim to have known it all along, Ofir and Mazursky (1997) hold that this may be because their acknowledged surprise triggered ''special processing'' that offset or reversed the process that normally produces hindsight bias. Indeed, other researchers have suggested that additional, qualitatively different, processing occurs when people are surprised by an outcome. For example, the experience of surprise has been posited to cause people to ''rethink the judgement'' (Wasserman et al., 1991, p. 35), to ''counteract the (hindsight) tendency to integrate the feedback into the subjects' general knowledge structures'' (Hawkins & Hastie, 1990, p. 315), or to ''serve as a cue to the outcomes' unexpectedness (which) reduces, even eliminates, traditional hindsight effects'' (Louie, 1999, p. 30).

Although such a special mechanism is possible, the present model indicates that it may be unnecessary. Sherman, Cialdini, Schwartzman, and Reynolds (1985) found that participants who tried to imagine a difficult-to-imagine outcome perceived the outcome as *less* likely than if they had not tried to imagine it at all. Apparently, trying to imagine an event, but failing to do so, causes that event's perceived likelihood to decrease. This may be because difficult-to-imagine, or ''unintegratable'' information tends to be forgotten or de-emphasised (Fischhoff & Beyth, 1975). Failure to integrate factors predictive of the known outcome likely produces the same effects as the established debiasing strategy of ''considering the opposite'' (Arkes et al., 1988; Davies, 1987; Slovic & Fischhoff, 1977). Thus, the reduction of the hindsight bias may be caused by a failure of the sense-making process, rather than by an additional mechanism that overrides it. The negative correlation between ''resultant'' surprise and hindsight bias magnitude in Experiments 2 and 4 seems consistent with this idea. Of course, Hawkins and Hastie (1990) may have anticipated this argument when they asked ''to what extent is surprise a by-product of the subjective difficulty experienced while integrating outcome feedback . . . ?'' (p. 324). Two predictions of the sense-making model were supported in this research, namely that relatively unexpected outcomes that are easy to make sense of will produce hindsight biases and that unexpected outcomes that are difficult to make sense of will not lead to hindsight biases.

Interestingly, cultural factors may moderate the need to explain unexpected events. Choi and Nisbett (2000) recently reported that the more surprising their participants found the results of a psychological research study, the less hindsight bias they exhibited. From the model, we assume that they refer to resultant surprise. More interesting, however, was the fact that Korean students in their study consistently reported less surprise (and more hindsight bias) at the unexpected outcomes than did American students. Choi and Nisbett contend that people from Asian cultures are less bothered by inconsistency, and thus do not spontaneously consider alternative outcomes after receiving an unexpected outcome in the way that Westerners would.

Because no direct measures of ''sensemaking'' were taken in this research, of course, the findings do not offer unequivocal evidence that sense-making success is the key determinant of hindsight bias. However, they do suggest that future research should consider this possibility. One direction this research might take is to observe brain activity (e.g., PET, fMRI) as participants receive unexpected outcomes that are either difficult or easy to understand. Difficult outcomes may force participants to ''recruit'' multiple areas of the cortex in an attempt to find a solution. A large amount of brain activity might be a sign that something is amiss, an indication of sense-making difficulty. One way to create such a situation might be to use Sherman et al.'s (1985) approach by asking participants to think of an overly large number of reasons for an outcome. If they attempt to do so, but fail, they may show increased or

altered brain activity, and report likelihood estimates representing little or reverse hindsight bias.

The motivational perspective: Defensive processing

A secondary goal of this paper was to examine the limits of the defensive-processing hypothesis (Louie et al., 2000; Mark & Mellor, 1991). This motivationally based idea also suggests that there are times when the hindsight bias will not occur. Here, however, the idea is that people can be motivated to ignore negative and self-relevant outcomes that threaten their self-image. Indeed, two studies in this special issue of *Memory* present evidence for this claim (Mark et al., 2003-this issue, and Renner, 2003-this issue). Data presented in this paper, however, imply important limitations for this approach.

Arguments for a defensive-processing mechanism have intuitive appeal. However, because negative outcomes—particularly those that are self-relevant—are often unexpected, care should be taken to also consider the effects of outcome expectations (Falk, 1989; Taylor & Brown, 1988; Weiner, 1985). The data presented here suggest that outcomes that were incongruent with expectations produced significant hindsight biases, regardless of self-relevance. Put differently, the stockbroker and Clinton supporter described in the introduction may very well exhibit hindsight bias in response to their upsetting outcomes, provided those outcomes are at least somewhat unexpected and some sense is able to be made of them.

In Experiment 1, for example, fans who expected their basketball team to win (and were correct) showed no hindsight bias, but fans of the losing team, who expected their team to win (but were incorrect) did show the bias. Although sports outcomes can have relevance for fans' self-esteem (Cialdini et al., 1976), the fans could not claim responsibility for the outcome. Thus, feelings of culpability may be required before a defensive-processing mechanism is activated.

In Experiment 2, participants could reasonably be said to feel responsible for their performance on a cognitive abilities test. Here, low self-esteem individuals did not exhibit hindsight bias after negative test feedback, but did exhibit the bias after positive feedback. This would seem to support the defensive-processing hypothesis. However, participants with high self-esteem were found to exhibit the opposite. They did not show the bias for positive feedback, but showed a sizeable bias for negative feedback. Unless there is a reason that defensive processing would occur only for those with low self-esteem, this is somewhat problematic for the defensive-processing hypothesis. In fact, previous research has found that high rather than low self-esteem individuals are most likely to utilise self-esteem-defensive mechanisms (Agostinelli et al., 1992; Brown & Gallagher, 1992).

Interestingly, Wegner's (1997) ironic mental processing model suggests that attempts to ignore threatening outcomes may backfire. Wegner found that trying not to think about a ''white bear'' produced significantly more thoughts about it than if no such attempt was made. In the defensive-processing model, a threatening outcome is essentially the white bear to be ignored. Indeed, anecdotal evidence suggests that threatening outcomes may actually produce greater hindsight bias. For example, feelings of guilt and self-blame can arise after concluding that we ''should have known'' and thus been able to prevent a tragic event from occurring to a loved one. Tykocinski (2001) has even suggested that we might feel *better* following an unpleasant outcome if that outcome seems relatively inevitable in hindsight. That is, it may be comforting to realise that nothing could have been done to avoid the outcome. Clearly, defensive-processing hypotheses present fascinating possibilities. Future research, however, will need to clearly define those situations in which defensive strategies are expected to mitigate the hindsight bias, rather than exacerbate it.

REFERENCES

Agostinelli, G., Sherman, S.J., Presson, C.C., & Chassin, L. (1992). Self-protection and self-enhancement biases in estimates of population prevalence. *Personality and Social Psychology Bulletin*, *18*, 631–642.

Alicke, M.D., Davis, T.L., & Pezzo, M.V. (1994). *A posteriori* revision of *a priori* decision criteria. *Social Cognition*, *12*, 281–308.

Arkes, H.R. (1988). Commentary on the article by Verplanken & Pieters. *Journal of Behavioral Decision Making*, *1*, 146.

Arkes, H.R., Faust, D., Guilmette, T.J., & Hart, K. (1988). Eliminating the hindsight bias. *Journal of Applied Psychology*, *73*, 305–307.

Brown, J.D., & Gallagher, F.H. (1992). Coming to terms with failure: Private self-enhancement and public self-effacement. *Journal of Experimental Social Psychology*, *28*, 3–22.

Cannon, C.K., & Quinsey, V.L. (1995). The likelihood of violent behaviour: Predictions, postdictions, and hindsight bias. *Canadian Journal of Behavioural Science*, *27*, 92–106.

Choi, I., & Nisbett, R.E. (2000). Cultural psychology of surprise: Holistic theories and recognition of contradiction. *Journal of Personality and Social Psychology*, *79*, 890–905.

Christensen-Szalanski, J.J.J., & Willham, C.F. (1991). The hindsight bias: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *48*, 147–168.

Cialdini, R.B., Borden, R.J., Thorne, A., Walker, M.R., Freeman, S., & Sloan, L.R. (1976). Basking in reflected glory: Three (football) field studies. *Journal of Personality and Social Psychology*, *34*, 366–375.

Davies, M.F. (1987). Reduction of hindsight bias by restoration of foresight perspective. *Organizational Behavior and Human Decision Processes*, *40*, 50–68.

Falk, R. (1989). Judgment of coincidences: Mine versus yours. *American Journal of Psychology*, *102*, 477–493.

Fischhoff, B. (1975). Hindsight $\neq$ Foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 288–299.

Fischhoff, B. (1977). Perceived informativeness of facts. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 349–358.

Fischhoff, B., & Beyth, R. (1975). ''I knew it would happen . . .'' Organizational Behavior and Human Performance, *13*, 1–16.

Fleming, J.S., & Courtney, B.E. (1984). The dimensionality of self-esteem: II. Hierarchical facet model for revised measurement scales. *Journal of Personality and Social Psychology*, *46*, 404–421.

Guerin, B. (1982). Salience and hindsight biases in judgements of world events. *Psychological Reports*, *50*, 411–414.

Haslam, N., & Jayasinghe, N. (1995). Negative affect and hindsight bias, *Journal of Behavioral Decision Making*, *8*, 127–135.

Hastie, R. (1984). Causes and effects of causal attribution. *Journal of Personality and Social Psychology*, *46*, 44–56.

Hawkins, S., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, *167*, 311–327.

Hoch, S.J., & Loewenstein, G.F. (1989). Outcome feedback: Hindsight and information. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *15*, 605–619.

Lau, R.R., & Russell, D. (1980). Attributions in the sports pages: A field test of current hypotheses in attribution research. *Journal of Personality and Social Psychology*, *39*, 29–38.

Leary, M. R. (1981). The distorted nature of hindsight. *Journal of Social Psychology*, *115*, 25–29.

Louie, T.A. (1999). Decision makers' hindsight bias after receiving favorable and unfavorable feedback. *Journal of Applied Psychology*, *84*, 29–41.

Louie, T.A., Curren, M.T., & Harich, K.R. (2000). ''I knew we would win'': Hindsight bias for favorable and unfavorable team decision outcomes. *Journal of Applied Psychology*, *85*, 264–272.

Mark, M.M., & Mellor, S. (1991). Effect of self-relevance of an event on hindsight bias: The foreseeability of a layoff. *Journal of Applied Psychology*, *76*, 569–577.

Mark, M.M., & Mellor, S. (1994). ''We don't expect it happened'': On Mazursky and Ofir's (1990) purported reversal of the hindsight bias. *Organizational Behavior and Human Decision Processes*, *57*, 247–252.

Mark, M.M., Boburka, R. R., Eyssell, K.M., Cohen, L.L., & Mellor, S. (2003). ''I couldn't have seen it coming'': The impact of negative self-relevant outcomes on retrospections about forseeability. *Memory*, *11*, 443–454.

Markman, K.D., & Tetlock, P.E. (2000). ''I couldn't have known'': Accountability, foreseeability and counterfactual denials of responsibility. *British Journal of Social Psychology*, *39*, 313–325.

Mazursky, D., & Ofir, C. (1990). I could never have expected it to happen: The reversal of the hindsight bias. *Organizational Behavior and Human Decision Processes*, *46*, 20–33.

Mazursky, D., & Ofir, C. (1996). ''I knew it all along'' under all conditions? Or possibly ''I could not have expected it to happen'' under some conditions. *Organizational Behavior and Human Decision Processes*, *66*, 237–240.

Menec, V.H., & Weiner, B. (2000). Observers' reactions to genetic testing: The role of hindsight bias and judgments of responsibility. *Journal of Applied Social Psychology*, *30*, 1670–1690.

Nario, M.R., & Branscombe, N.R. (1995). Comparison processes in hindsight and causal attribution. *Personality and Social Psychology Bulletin*, *21*, 1244–1255.

Ofir, C., & Mazursky, D. (1997). Does a surprising outcome reinforce or reverse the hindsight bias? *Organizational Behavior and Human Decision Processes*, *69*, 51–57.

Olson, J.M., Roese, N.J., & Zanna, M.P. (1996). Expectancies. In E.T. Higgins & A.W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles*. New York: Guilford Press.

Pezzo, M.V. (1996). Removing the hindsight bias: A test of the motivated processing hypothesis. (Ohio University, 1996). *Dissertation Abstracts International*, *57*, 6606.

Pohl, R.F. (1998). The effects of feedback source and plausibility of hindsight bias. *European Journal of Cognitive Psychology*, *10*, 191–212.

Pyszczynski, T.A., & Greenberg, J. (1981). Role of disconfirmed expectations in the instigation of attributional processing. *Journal of Personality and Social Psychology*, *40*, 31–38.

Raven, J., Raven, J.C., & Court, J.H. (1995). *Manual for Raven's Progressive Matrices and vocabulary Scales*. Oxford: Oxford University Press.

Renner, B. (2003). Hindsight bias after receiving self-relevant health risk information: A motivation perspective. *Memory*, *11*, 455–472.

Roese, N.J., & Maniar, S.D. (1997). Perceptions of purple: Counterfactual and hindsight judgments at

Northwestern Wildcats football games. *Personality and Social Psychology Bulletin*, *23*, 1245–1253.

Roese, N.J., & Olson, J.M. (1996). Counterfactuals, causal attributions, and hindsight bias: A conceptual integration. *Journal of Experimental Social Psychology*, *32*, 197–227.

Sanna, L.J., Schwarz, N., & Stocker, S.L. (2002). When debiasing backfires: Accessible content and accessibility experiences in debiasing hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 497–502

Sanna, L.J., & Turley, K.J. (1996). Antecedents to spontaneous counterfactual thinking: Effects of expectancy violation and outcome valence. *Personality and Social Psychology Bulletin*, *22*, 906–919.

Schkade, D.A., & Kilbourne, L.M. (1991). Expectation–outcome consistency and hindsight bias. *Organizational Behavior and Human Decision Processes*, *49*, 105–123.

Schwarz, N. (1998). Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality and Social Psychology Review*, *2*, 87–99.

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.

Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 544–551.

Snyder, C.R., & Higgins, R.L. (1988). Excuses: Their effective role in the negotiation of reality. *Psychological Bulletin*, *104*, 23–35.

Tan, H.-T., & Lipe, M.G. (1997). Outcome effects: The impact of decision process and outcome controllability. *Journal of Behavioral Decision Making*, *10*, 315–325.

Taylor, S.E., & Brown, J.D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193–210.

Tykocinski, O.E. (2001). I never had a chance: Using hindsight tactics to mitigate disappointments. *Personality and Social Psychology Bulletin*, *27*, 376–382.

Verplanken, B., & Pieters, R.G.M. (1988). Individual differences in reverse hindsight bias: I never thought something like Chernobyl would happen. Did I? *Journal of Behavioral Decision Making*, *1*, 137–147.

Wasserman, D., Lempert, R.O., & Hastie, R. (1991). Hindsight and causality. *Personality and Social Psychology Bulletin*, *17*, 30–35.

Wegner, D.M. (1997). When the antidote is the poison: Ironic mental control processes. *Psychological Science*, *8*, 148–150.

Weiner, B. (1985). ''Spontaneous'' causal thinking. *Psychological Bulletin*, *97*, 74–84.

Winman, A. (1997). The importance of item selection in ''knew-it-all-along'' studies of general knowledge. *Scandanavian Journal of Psychology*, *38*, 63–72.

Zakay, D. (1984). The influence of perceived event's controllability on its subjective occurrence probability. *Psychological Record*, *34*, 233–240.