APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. **22**: 491–506 (2008)
Published online 27 June 2007 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/acp.1377

The Effects of Disappointment on Hindsight Bias for Real-World Outcomes

MARK V. PEZZO[1]* and JASON. W. BECKSTEAD[2]

[1] *University of South Florida, St. Petersburg, USA*
[2] *College of Nursing, University of South Florida, USA*

SUMMARY

Examined the effects of motivated processing on hindsight bias. One hundred fifty three college students estimated and later recalled the likelihood that 30 self-relevant events would occur during the next 2 months. Multi-level modelling was used to determine the (within-subject) effects of expectations, event valence and event controllability on hindsight bias and the extent to which these effects were moderated by participants' need for cognition (NFC) scores (between-subjects). For events that actually occurred, we found support for defensive processing in that the bias was smaller for negative events. Also, for events that actually occurred, those judged as more controllable produced a larger bias. Neither valence nor controllability had any effect on the size of the bias for events that did not occur. The size of the bias for occurrences did not differ significantly from that for non-occurrences. Finally, NFC did not moderate the effects of valence or controllability, nor did it directly affect the size of the bias. Copyright © 2007 John Wiley & Sons, Ltd.

Wish I didn't know now what I didn't know then.

—Bob Seger

Failure and disappointment are inevitable. New businesses, political campaigns and romantic pursuits fail more often than they succeed. When they do, we usually try to make sense of what went wrong (Wilson, Gilbert, & Centerbar, 2003). Disappointed sports fans and ex-lovers alike struggle to understand in hindsight what they didn't know in foresight. Whether or not we learn from such failures may depend on the degree to which we experience hindsight bias (Pezzo & Pezzo, 2007). Also referred to as the knew-it-all-along-effect, the hindsight bias occurs when we overestimate how predictable an outcome was in foresight. This phenomenon has been studied extensively (Christensen-Szalanski & Willham, 1991; Guilbault, Bryant, Brockway, & Posavac, 2004; Hawkins & Hastie, 1990) since first proposed by Fischhoff (1975). Current research has primarily focused on two issues—the degree to which surprising outcomes enhance, reduce, or reverse the bias (e.g. Ofir & Mazursky, 1997; Pezzo, 2003; Schkade & Kilbourne, 1991), and the degree to

*Correspondence to: Mark V. Pezzo, Psychological Sciences, USFSP, St. Petersburg, DAV 258, 140 S. 7th Ave, St. Petersburg, FL 33701, USA. E-mail: pezzo@stpt.usf.edu

which self-relevance of the outcome plays a role in enhancing, reducing, or reversing the bias. This paper will focus primarily on the second issue.

Very few hindsight bias studies have used self-relevant real-world outcomes. Of the 95 published and unpublished papers reported in Guilbault et al.'s (2004) recent meta-analysis, only 21 used real-world outcomes.[1] Among those, most used outcomes that either had little relevance to the *self* (e.g. the outcome of President Nixon's trip to China) or failed to measure the meaning of the outcome to participants. For example, Pennington, Rutter, McKenna, and Morley (1980) examined women's predictions of whether or not they were pregnant in foresight or hindsight, but did not ask whether or not the women *wanted* to become pregnant. As discussed in the next section, we might expect the size of the hindsight bias to differ for people who are disappointed or threatened by the outcome than for those who are not.

DEFENSIVE PROCESSING OF NEGATIVE OUTCOMES

Some researchers suggest that negative self-relevant outcomes may be less susceptible to the hindsight bias. Mark and Mellor's (1991) seminal research found that laid-off employees reported their layoff as less foreseeable than did either retained employees or community members. Because this study did not also record foresight judgements, however, we cannot be sure that the findings resulted from a self-protective motive. In a second, experimentally controlled study (Mark, Boburka, Eyssell, Cohen, & Mellor, 2003) players in a simulated stock market game lost money and judged the outcome less foreseeable than did their opponent or an observer (for whom the choice had no self-relevance). Self-relevance did not affect the amount of hindsight bias, however, for positive outcomes in this study.

Louie (1999, Study 1) also examined hindsight bias in a simulated stock purchase. Participants receiving an upsetting outcome (i.e. purchased stock that lost value or did not purchase stock that increased in value) showed no significant hindsight bias, while those receiving a positive outcome did exhibit the bias. In a second study (Louie, Curren, & Harich, 2000) MBA students exhibited hindsight bias following the poor performance of a competing team or their own good performance, but no bias for their own failure or a competing team's success. Stahlberg and Schwarz (1999) provided similar evidence for defensive processing following negative (simulated) job interviews.

To date, only three studies have examined defensive processing in a real-world setting, using a true control condition and these have obtained mixed results. Hölzl, Kirchler, and Rodler (2002) found that supporters of the Euro showed less hindsight bias when its value dropped than when it increased. Renner (2003) asked participants to predict their approximate cholesterol level (e.g. high, medium, low) approximately 40 minutes before being told their actual score, immediately after receiving their score, and then again, approximately 5 weeks later, whereas those receiving unexpectedly positive outcomes did not exhibit hindsight bias, those who received unexpected negative results showed hindsight bias at first, but reverse hindsight bias after 5 weeks. That is, they recalled a lower prediction than the already low one they made in foresight! Renner suggests that hindsight bias might reflect an attempt to regain a sense of control and self-efficacy in the face of an

[1] We thank Rebecca Guilbault and Fred Bryant for supplying the data set from their (2004) meta-analysis.

DOI: 10.1002/acp

uncontrollable threat. The eventual emergence of a reverse hindsight bias may occur only after the feelings of threat dissipate and people's goals shift to one of avoiding blame.

In a third study (Haslam & Jayesinghe, 1995), however, college students who were disappointed by an exam grade exhibited typical hindsight bias. Students who received a better grade than expected showed reverse hindsight bias, recalling an even lower predicted score! Haslam and Jayasinghe speculated that students may have been attempting to increase their positive affect by magnifying the idea that they did better than expected. Although this finding contradicts both Renner (2003) and Hölzl et al. (2002), it seems consistent with a second proposed mechanism called retroactive pessimism.

RETROACTIVE PESSIMISM: A DIFFERENT TYPE OF DEFENSIVE PROCESSING

In contrast to previous findings, Tykocinski (2001) reports a 'retroactive pessimism' effect in which hindsight bias is *greater* following negative outcomes, presumably because 'an inescapable failure might be easier to digest than a failure that could have been easily avoided' (p. 381). In her Experiment 1, people rated the likelihood of a favourable outcome after imagining that a disappointing outcome actually occurred. Larger hindsight bias occurred for more disappointing outcomes, but only for participants scoring high in the desire for control (Burger & Cooper, 1979). In Experiment 2, people who were disappointed with the result of an election showed greater hindsight bias than those pleased with the election. However, this effect was only obtained for likelihood ratings of the winning candidate, not the losing candidate. Tykocinski argued that because the losing candidate quickly receded into the background, his loss was somewhat akin to a non-occurrence, typically found to produce smaller hindsight bias. Tykocinski, Pick, and Kedmi (2002, Experiment 1) found that fans of a losing soccer team showed greater hindsight bias than did fans of the winning team (see also Pezzo, 2003, Experiment 1). Most important, the magnitude of the bias was related to the degree of disappointment expressed by losing fans. Finally, Tykocinski et al. (Experiments 2 and 3) found that participants who imagined that they missed a deadline to apply for a college stipend exhibited greater hindsight bias when the stipend was large than when it was small. It is important to note that in none of the studies reported by Tykocinski and her colleagues was the participant *responsible* for the outcome.

Some studies have not found clear evidence for either form of defensive processing. For example, Pezzo (2003, Experiment 2) found that participants with low academic self-esteem showed hindsight bias for positive, but not negative, feedback on a 'cognitive abilities' test. High self-esteem participants showed the opposite—greater hindsight bias for negative than for positive feedback. Although this may be taken as defensive processing for those low in self-esteem and retroactive pessimism for those high in self-esteem, a more parsimonious explanation may be that unexpected outcomes produce the greatest hindsight regardless of valence (Roese & Olson, 1996; Schkade & Kilbourne, 1991).

OUTCOME CONTROLLABILITY

The defensive processing approach argues that we can reduce our sense of culpability by denying the foreseeability of a negative event (Mark et al., 2003). Events that are not under

DOI: 10.1002/acp

one's control should not require defensive processing because they do not threaten the self. According to a recent theoretical review by Pezzo and Pezzo (2007), one potential difference between defensive processing and retroactive pessimism studies is that the latter typically uses outcomes that were produced by relatively uncontrollable events. Indeed, Tykocinski and Steinberg (2005) found that retroactive pessimism only occurred when an imagined negative outcome resulted from an uncontrollable, rather than a controllable event. Others, however, report that hindsight bias is less likely to occur for uncontrollable events (e.g. Tan & Lipe, 1997; Wasserman, Lempert, & Hastie, 1991), presumably, because such events have fewer causal antecedents. Thus, the present study will measure outcome controllability to determine if it has any direct (cognitive) effects on the hindsight bias, and whether it moderates the (motivational) effects of outcome valence.

NEED FOR COGNITION

Finally, because hindsight bias is most commonly thought to be caused by an underlying sense-making mechanism, it may be related to the *extent* to which one prefers to engage in effortful thought. Thus, although a motivational mechanism may determine the *direction* of the bias, one's need for cognition (NFC, Cacioppo, Petty, & Kao, 1984) may moderate the *magnitude* of the bias. Very few studies have examined the effects of NFC on hindsight bias. One exception, Verplanken and Pieters (1988), reported that high NFC participants showed the least hindsight bias for a real-world event (Chernobyl nuclear accident), whereas medium and low NFC participants actually showed a reverse hindsight bias. Although Arkes (1988) believes that their interpretation of the bias as reversed is mistaken, this does not discount their claims that NFC moderated the bias. Musch (2003) however, found no effect of NFC when using general knowledge questions. To our knowledge, no other hindsight bias researchers have used the NFC scale, prompting its inclusion in the present study.

OVERVIEW

The use of self-relevant real-world outcomes is rare in hindsight bias research, in part, because such outcomes cannot be randomly assigned. However, the benefits of using real-world events may sometimes outweigh the failure to randomly assign outcomes, particularly when random assignment is ethically questionable or otherwise impossible (Renner, 2003). We asked participants to make foresight and hindsight predictions for a large number of real-world events ($n = 30$). In doing so, we hoped to increase the chances that each participant could provide data on a wide *range* of events that varied in valence, likelihood of occurrence and controllability. This approach exemplifies Brunswik's (1956) representative design. We used a multi-level modelling approach that allows for both within-subject and between-subject analyses. A within-subjects analysis tested the hypothesis that 'unhappy' outcomes would produce less hindsight bias than 'happy' outcomes, particularly when they are controllable. We also explore the possibility, using a between-subjects analysis that NFC may directly affect the size of the hindsight bias and also that it may moderate the effects of valence and controllability on the bias. The data resulting from our design are unbalanced in that participants neither respond to the same

DOI: 10.1002/acp

number, nor even the same subset of events. While this situation presents problems for an OLS regression approach, such data structures may be analysed quite easily by a multilevel model.

METHOD

Participants

One hundred fifty three Introductory Psychology students[2] attending a selective private university received course credit, and were entered in a raffle to win $25 if they completed both phases of the experiment.

Procedure

At the beginning of the Fall semester, participants completed the 18-item NFC scale (Cacioppo et al., 1984). Three to four weeks preceding the Thanksgiving Holiday, we asked participants to indicate the likelihood that each of 30 events[3] would occur between 'now and the beginning of next semester'. Thirteen of these events were positive and 17 were negative (see Table 1). Participants provided a likelihood value between 0% and 100% for each event. If an event was not applicable (e.g. participant who does not drive asked likelihood of getting speeding ticket) the student was instructed to write 'NA', rather than 0%.

Approximately 8–9 weeks after the first session, participants were telephoned and asked to return for a short 15-minute session. One hundred twenty six participants (82%) returned for the second session. Here, they were asked to indicate whether or not each of the events had actually occurred, and then to recall their original likelihood estimate, using the same 0–100% scale as before.

A separate sample of 52 introductory psychology students were asked to rate the controllability of each of the 30 events, using a 6-point scale ranging from 'extremely *un*controllable' to 'extremely controllable'.

ANALYSIS PROCEDURES

Because each participant responded to a different number of events, our nested and unbalanced design violates independence assumptions. Although one could aggregate the responses of each participant and analyse mean or median scores, this approach reduces statistical power and ignores within-subject variability, which may account for a large proportion of total variability (Pezzo & Beckstead, 2006). Instead, we employed multilevel modelling (Bryk & Raudenbush, 2002). Besides making efficient use of all data without violating non-independence assumptions, the multi-level approach typically provides more stable regression coefficients and more accurate estimates of their standard errors than the

[2]Data on participant sex and age were not collected, but these variables have typically not been shown to affect the hindsight bias.
[3]We originally started with 51 events, but were concerned that some might not have universally agreed upon valence. As a manipulation check, we asked 47 additional students to indicate whether each of the events should be considered positive or negative if it occurred. Twenty-one events produced marked disagreement and were dropped. Of the remaining 30 events, the average agreement was 97.4%, with only 5 events (3 negative, 2 positive) receiving below 98% agreement (94%, 91%, 91%, 85% and 80%).

DOI: 10.1002/acp

Table 1. Mean likelihood predictions, percent actually happening and controllability ratings for positive and negative events

Event	n[a]	Original judgement	*SD*	% actually happened	Controllability[b]	*SD*
Positive events ($n = 13$)						
Have a good Thanksgiving break	125	86.8	14.3	92.0	4.58	0.9
Have a good Christmas break	125	90.7	12.6	91.2	4.73	0.7
Make a good friend	124	67.6	24.9	88.7	4.31	1.3
Hear from an old friend	125	76.8	24.0	84.0	2.67	1.3
Do really well on an important exam	125	64.0	21.2	81.6	5.25	0.6
Have a good road trip	122	72.0	26.0	76.2	5.00	0.8
Favourite team has great season	104	65.9	28.8	74.0	1.25	0.8
Have a good romantic ‘hook up’	118	39.7	31.9	46.6	4.54	1.1
Get asked out by someone you like	118	39.4	29.8	44.9	3.73	0.8
Get nominated for an honour/award	124	21.7	20.6	17.7	4.17	1.2
Tell someone you love them	117	22.1	26.0	12.8	4.83	1.4
Get a new car	121	6.2	16.7	5.8	4.12	1.6
Win a contest/lottery	123	12.4	15.6	4.9	1.46	0.8
Negative events ($n = 17$)						
Favourite team loses a big game	108	63.4	25.0	86.1	1.10	0.6
Get into a fight with your parents	124	37.5	29.7	42.7	4.13	0.9
Get really sick from drinking	120	20.1	30.7	33.3	5.40	1.0
Get a parking/speeding ticket	116	28.9	21.8	32.8	5.31	0.9
Get very sick	124	27.5	22.1	30.6	2.52	1.2
Forget to pay a bill	121	19.2	21.9	28.1	5.02	1.1
Get a ‘D’ or ‘F’ on an exam	124	36.1	31.0	27.4	5.13	0.7
Lose your keys or wallet	124	32.1	25.4	25.8	4.13	1.3
Have a bad romantic ‘hook up’	116	26.4	27.0	24.1	4.25	1.2
Be discriminated against	124	29.5	27.4	21.8	1.79	1.0
Get a ‘D’ or ‘F’ in a class	122	20.8	27.4	13.1	5.06	0.9
Go to the emergency room	124	18.0	21.2	8.9	2.94	1.2
Lock your keys in your car	107	20.2	21.4	8.4	4.54	1.4
Partner cheats on you	73	14.5	20.2	8.2	2.06	1.0
Be involved in an auto accident	122	22.3	16.7	7.4	3.08	1.2
Your pet dies	83	25.9	26.3	7.2	1.46	0.6
Fake ID taken from you	40	15.4	21.4	2.5	4.33	1.3

[a]Number of participants responding to each event (total events = 3443).
[b]Controllability ratings provided by 52 additional students not participating in experiment.

traditional Ordinary Least-Squares method.[4] A key advantage to multi-level modelling is that events for which fewer participants responded are not weighed as heavily in the analysis (see Nezlek, 2001 for further discussion).

Hindsight bias magnitude

We defined *bias* as the difference between the original likelihood judgement (OJ) and recall of that judgement (ROJ) for a given event (ROJ−OJ). Positive bias scores indicate that the

[4]More specifically, the regression coefficients obtained in our within-subjects (level-1) analysis become dependent variables in regressions in our between-subjects (level-2) analysis. The degree of dependence among observations (i.e. those provided by the same participant) is estimated over all participants and taken into account as the standard errors are calculated for regression coefficients, thus reducing the bias introduced by violating the independence assumption.

DOI: 10.1002/acp

recalled estimate was greater than the original (ROJ > OJ), and represents hindsight bias for an event that occurred. However, it indicates a reverse hindsight bias for an event that did not occur. Negative bias scores (ROJ < OJ) indicate the opposite—hindsight bias for an event that did *not* occur, and reverse hindsight bias for an event that *did* occur.

Dichotomous predictors

Two variables *valence* (V) and *happened* (H) are dichotomous. Contrast coding was used to indicate the valence of each event (positive = +0.5; negative = −0.5), and whether or not each event actually happened (yes = +0.5; no = −0.5). The regression coefficients obtained for these two variables (see Table 2) indicate the difference in effect on hindsight bias between the two levels of each variable. Multiplying each coefficient by the respective contrast code yields the condition's treatment effect relative to the grand mean, as they are typically expressed in the ANOVA model.

Continuous predictors

Participants' *original judgement* (OJ) and event *controllability* scores (C) are both continuous variables; their coefficients (see Table 2) reflect unit change relative to their means (see Table 2).

Level-1 analyses

At level-1, the amount of hindsight bias for an event was modelled as a function of the participant's grand mean and (a) the *original judgement*, (b) whether or not the event actually *happened*, (c) the *valence* of the event, (d) the event's *controllability* and (e) three interactions (valence × happen; controllability × happen; valence × happen × controll- × controllability). Because hindsight bias will have a different sign depending on whether or not the outcome actually occurs, all interactions include the 'happen' term for proper interpretation. Further, note that the coefficients for event *valence* and *controllability* are not theoretically meaningful by themselves. For example, although a positive event that happens is a 'good' thing, a positive event that does not happen is not. Nevertheless, the

Table 2. Descriptive statistics for predictor variables and recalled judgement

Variable	*M*	*SD*	Interpretation
Original judgement	38.63	33.63	Many events were not expected to occur
Happened	−0.11	0.49	More events did not happen than did
Valence	−0.04	0.50	Slightly more of the events were negative[a]
Controllability	3.81	1.32	Events were moderately controllable overall
Recalled judgement	36.51	29.56	

Note: When determining regression coefficient for a given predictor, means in this table are used as default values for all remaining predictors.
[a]A negative event that does not occur is a 'positive outcome'.

DOI: 10.1002/acp

lower-order predictors of valence and controllability were included in the model to ensure that the higher order interactions would be accurately estimated (Aiken & West, 1991).

Level-2 analyses

All preceding variables were modelled as within-subjects predictors. NFC is an individual difference (between-subjects) variable, and was modelled at level-2 to determine if it moderates any of the level-1 effects.

Centring

At level-1, the original probability judgement (OJ) was centred at 50%. At level-2 each participant's grand mean was modelled as a function of the population mean and a random term reflecting between-subject variance. All main effect and interaction terms were grand-mean centred.

Regression to the mean

Because recalled estimates were imperfectly correlated with predictions ($r = 0.69$), they are subject to regression to the mean effects (Darlington, 1990), which can be mistaken for a hindsight bias (Bryant & Brockway, 1997).[5] Inclusion of the original judgement term (OJ) in the regression analysis ensures that all other effects are free of regression to mean.

RESULTS

Preliminary results

Table 1 indicates that most of our 126 participants responded to most events ($M = 115$, $SD = 18.1$). Some events, however, were clearly less applicable than others (e.g. only $n = 40$ responded to 'Fake ID taken from you'). Table 1 also shows the mean original probability judgement for each event, the percentage of applicable people for whom the event actually happened, and the mean controllability rating provided by a separate sample of 52 participants. Mean original judgements ranged from very low ($M = 6.21\%$; get a new car) to quite high ($M = 90.7\%$; have a good Christmas break). Controllability ratings also ranged from very low ($M = 1.10$; favourite team loses a big game) to quite high ($M = 5.40$; get sick from drinking). Thus, our sample of real-world events captured a wide range of judged likelihoods and perceived controllability. The total number of events responded to was 3443.

[5]The expected amount of regression to the mean is $100(1–r)$, and is equal to $(100(1–0.69) = 31\%$ in the present study. Thus, if the mean original judgement were, for example, 50% for some event, but a particular participant gave an original judgement of only 20%, we would expect his *recalled* judgement to be $0.31(50–20) = 9.3$ points closer to the mean, even if no hindsight bias occured.

DOI: 10.1002/acp

Table 3. Summary of within-subjects (level-1) analysis of bias in recall of event likelihood[a]

Source	Symbol	β	SE	p	η^2
Intercept		−2.05	0.67	0.003	0.056
Original judgement	OJ	−0.60	0.02	<0.001	0.88
Happened	H	20.30	1.21	<0.001	0.696
Outcome expectedness	OJ × H	0.07	0.03	0.023	0.042
Outcome valence	V × H	8.13	1.54	<0.001	0.185
Outcome controllability	C × H	1.89	0.51	<0.001	0.10
Three-way interaction	V × H × C	0.95	0.97	0.325 (n.s.)	0.008
Event valence[b]	V	2.39	0.84	0.005	—
Event controllability[b]	C	0.62	0.26	0.019	—
Two-way interaction[b]	V × C	0.69	0.44	0.114 (n.s.)	—

Note: $N = 3443$. All effect sizes calculated using df = 123.
[a]Bias = Recalled judgement–original judgement.
[b]Event valence, controllability and their interaction are not directly interpretable. For example, the effect of an event's valence without also considering whether the event actually happened is misleading. Interpretable effects can only be found via interactions between these variables and happened.

Level-1 predictors

Table 3 provides results of the level-1 analysis.[6] The intercept of −2.05 is the estimate of the grand mean of recall bias ($M_{\text{observed}} = -2.12$, $SD = 25.20$) and reflects the difference between the mean recalled and original judgements (36.51 and 38.63, respectively), averaged across all events. Interpretation of the coefficients for a given predictor is obtained by holding all other predictors constant at their mean values (see Table 2). Because all predictors are grandmean centred, any term held constant at its mean drops out of the equation.

Happened
As expected, whether or not an event happened significantly affected recall bias. Holding other predictors constant at their respective mean values, the model predicts a recall bias of $[-2.05 + 20.3(0.5 - (-0.11))] = 10.33$ for events that happened, and a predicted recall bias of $[-2.05 + 20.3\,(-0.5 - (-.11))] = -9.97$ for events that did not happen. Both values are significantly different than zero, $p\text{s} < 0.001$, indicating a typical hindsight bias in each case. The negative sign merely reflects the fact that recalled judgements decrease following non-occurrences. The absolute size of the hindsight bias for occurrences did not differ significantly from that of non-occurrences, $t\,(122) = 0.30$ $p = 0.76$.

Valence
As stated earlier, the coefficient representing the main effects of valence and controllability were included only to keep higher-order interactions accurate, and should not be interpreted alone. The *valence* × *happened* interaction, however, was significant, indicating different effects for 'happy' and 'unhappy' outcomes, and can be seen in Figure 1. First, consider the left side of Figure 1. Simple effects testing indicated that,

[6]Although not reported here, virtually the same coefficients were obtained when recalled judgements were used as the dependent measure instead of the difference score between recalled and original judgements. We chose the latter for its ease of interpretation with respect to hindsight bias.

DOI: 10.1002/acp

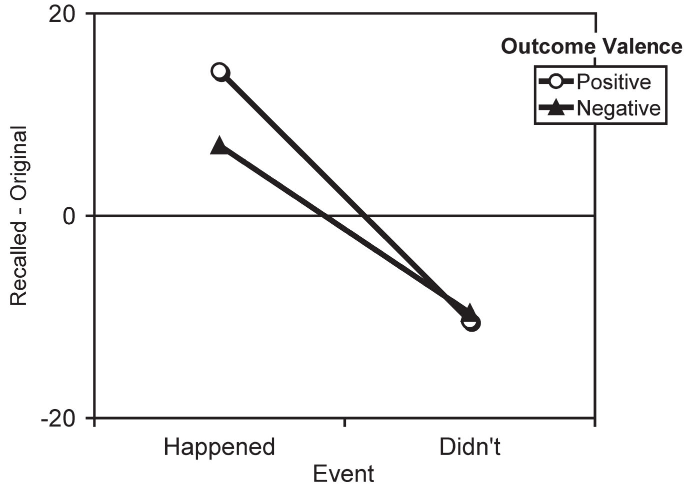

Figure 1. Effects of event valence on hindsight bias scores (recalled–original). Positive scores indicate hindsight bias for events that happened; negative scores indicate hindsight bias for events that did not happen. Hindsight bias is smaller for negative than for positive occurrences (left side of figure). For events that did not happen (right side), event valence had no effect

consistent with defensive processing, positive events that happened (happy outcomes) produced a significantly larger hindsight bias than negative events that happened (unhappy outcomes), t (122) = 5.13, $p < 0.001$, $\eta^2 = 0.18$. Now consider the right hand side of Figure 1. When the event did not happen, however, the size of the bias was unaffected by the event's valence, t (122) = 0.34, $p = 0.73$. The lack of a three-way interaction indicates that the controllability of an event did not qualify these findings.

Original judgement

In this paper, we refer to 'surprise' as how unexpected an event is. Events that are expected but do not occur, or that are not expected but do occur are both considered 'surprising'. [7,8] Figure 2 presents predicted hindsight bias scores as a function of the main effects of both original judgement, happened, and their interaction. Our hypothesis that surprising outcomes produce the greatest hindsight bias may be examined using simple main effects of *happened* at different levels of expectation. First, consider the top half of Figure 2. As predicted, the hindsight bias for unexpected events (OJ = 25%) was significantly larger when those events happened than when they did not, t (122) = 15.28, $p < 0.001$, $\eta^2 = 0.66$. Now consider the bottom half of Figure 2. Again as predicted, hindsight bias for expected events (OJ = 75%) was significantly larger when those events did not happen than when they did, t (122) = 18.21, $p < 0.001$, $\eta^2 = 0.73$. Finally, there was a significant H $\times$ OJ interaction, indicating that the slope for expected events (Figure 2, bottom) is very slightly steeper than for unexpected events (Figure 2, top). Because the size of this effect is very small and is overwhelmed by the simple main effects, care should be taken not to overstate its importance (Keppel & Wickens, 2004).

[7]The OJ values of 75% and 25% were chosen arbitrarily to represent expected and unexpected events, respectively.
[8]Although most researchers use 'surprise' and 'unexpectedness' interchangeably, Pezzo (2003) notes that unexpectedness is necessary *but not sufficient* for surprise. To experience surprise, one must fail to 'make sense' of an unexpected outcome. An unexpected outcome that one is able to make sense of will not produce the phenomenological experience of surprise.

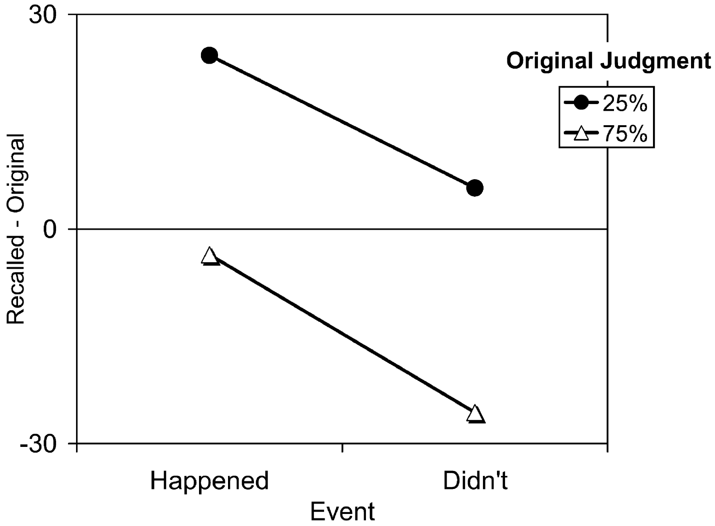

Figure 2. Effects of original judgement (expectations) on hindsight bias scores (recalled–original). For unexpected events (25%—top half of this figure) hindsight bias was greater when the event happened. For expected events (75%—bottom half of this figure) hindsight bias was greater when the event did *not* happen. In both cases, this represents a typical 'initial surprise' effect (Pezzo, 2003)

Controllability

As with event valence, the effect of event controllability on hindsight bias is not meaningful without also knowing whether or not that event occurred. Figure 3 shows that for events that actually happened, those rated high in controllability produced significantly greater hindsight bias than did those rated low. As indicated by the significant C $\times$ H interaction, however, when an event did not happen, controllability had no effect on the size

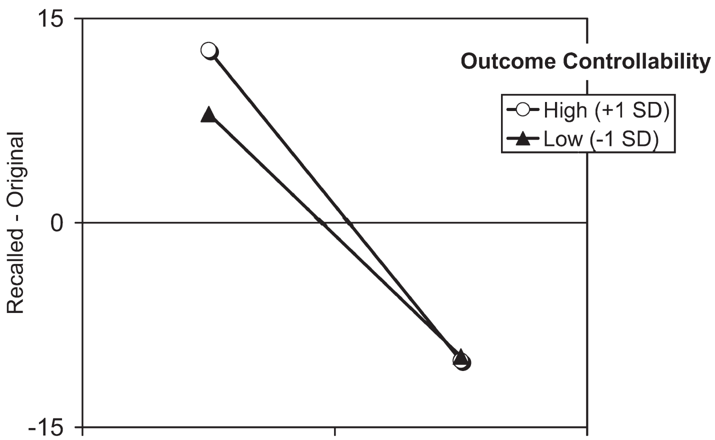

Figure 3. Effects of controllability on hindsight bias scores (recalled–original). For events that happened, hindsight bias is greater when event is more controllable. For events that did not happen, hindsight bias is unaffected by controllability. Negative scores indicate typical hindsight bias for events that did not happen

DOI: 10.1002/acp

of the hindsight bias. As noted above, the $V \times C \times H$ interaction, which indicates the extent to which defensive processing is moderated by event controllability, was not significant. Thus, the difference in hindsight bias for 'happy' vs. 'unhappy' outcomes was not affected by event controllability.

Level-2 analysis of need for cognition

Only three of the coefficients obtained at level-1 had a significant random component. The magnitude of the coefficients representing *original judgement* (OJ), *happened* (H) and their interaction ($OJ \times H$), all varied significantly among participants. This variability was modelled at level-2 to determine if NFC scores could account for any portion of it. Results indicated, however, that none of these three cross-level interactions were significant (ps > 0.60). Thus, NFC neither moderates the overall hindsight bias, nor the effect of expectations on the hindsight bias.

DISCUSSION

Very few studies have documented a motivated reduction in hindsight bias for real-world events. Of those that have (Hölzl et al., 2002; Renner, 2003), both used only a single outcome. The present study used 3443 self-relevant outcomes in a real-world setting to examine the hindsight bias. Unlike many laboratory studies, in which hindsight judgements are often made within an hour of the initial prediction, we waited approximately 2 months before asking people to recall their predictions. Because this study used as many as 30 different events per person, it is less vulnerable to stimulus sampling concerns (Wells & Windschitl, 1999). A number of factors were found to affect the size of the hindsight bias (see Table 3 for effect sizes). Some results are consistent with past research. Other findings, particularly when an event did not occur, were inconsistent with prior research.

Basic hindsight and surprise effects

As can be seen in Figures 1, 2 and 3, a strong hindsight bias was found overall. When an event actually occurred, recalled likelihood estimates were larger than original estimates, and when an event did not occur, recalled estimates were smaller. Unlike some past research (Christensen-Szalanski & Willham, 1991; Guilbault et al., 2004), the magnitude of the hindsight bias for our real-world outcomes did not differ significantly for occurrences and non-occurrences.

Figure 2 also shows that surprising events—those that were expected but did not happen, or those that were unexpected but did happen—produced substantially larger hindsight bias than did unsurprising events (Pezzo, 2003; Schkade & Kilbourne, 1991). This was the largest effect in our study. Although Figure 2 only shows the size of the hindsight bias for two values of expectation (25% and 75%) the *original judgement* measure varies continuously from 0% to 100%.

Reverse hindsight

Although the results clearly indicate a hindsight bias overall, Figure 2 also seems to suggest a small amount of reverse hindsight for *un*surprising outcomes—expected events that

DOI: 10.1002/acp

happened and unexpected events that did not happen. Reverse hindsight bias is somewhat controversial (e.g. Arkes, 1988; Mark & Mellor, 1994), although it has been increasingly reported (e.g. Menec & Weiner, 2000; Renner, 2003; Winman, Juslin, & Bjorkman, 1998). Normally, unsurprising outcomes produce little sensemaking effort, and thus little or no hindsight bias (Pezzo, 2003). One possibility, though speculative, is that as an event 'unfolds', at least some information becomes available that predicts an outcome that did *not* occur. Thus, even though the expected outcome did occur, the information contrary to this outcome momentarily activates the sense-making process, and causes judgements to shift slightly in the direction of the opposite outcome. Future research should explore this idea further.

Motivational factors

The valence × happened interaction was our largest interaction, by a considerable amount, and provides support for the defensive processing hypothesis (Mark et al., 2003; Louie et al., 2000), albeit with an important qualification. As can be seen in Figure 1 (left side), even after controlling for initial expectations, negative events that actually happen produce less hindsight bias than do positive events that happen. When an event did not happen (Figure 1, right side), we found no difference in the size of the bias between positive and negative events. Although non-occurrences (e.g. did not get asked out on a date) are ubiquitous in the real world, this is the first study we know of that has studied their effects on defensive processing. Clearly, a positive event that does not occur can pose a self-relevant threat, and should, according to defensive processing, produce less hindsight bias. Future research should continue to examine the possibility that defensive processing may not be activated by threatening *non*-occurrences.

Controllability

Support for defensive processing would seem to imply a lack of support for retroactive pessimism (Tykocinski, 2001). Recall, however that Tykocinski argues that outcome controllability determines which form of motivational processing occurs. We measured outcome controllability with the logic that one should not feel the need to defend against negative outcomes that were outside their own control (Pezzo & Pezzo, 2007).[9] Thus, for very controllable events, we expected to find the most defensive processing. Further, based on the idea of retroactive pessimism (Tykocinski, 2001; Tykocinski & Steinberg, 2005), we expected increased hindsight bias for very uncontrollable events. The lack of a significant three-way (valence × happened × controllability) interaction, however, indicates that controllability did not moderate the effects of outcome valence in our study.

Thus, our data do not support either motivational mechanism for events that do not happen, but seem to support a rather strong version of the defensive processing hypothesis—in that defensive processing occurs even for uncontrollable events that one would not think should pose much threat to a person's self (Mark et al., 2003). This study was unable to provide any support for retroactive pessimism despite including outcomes that were rated very low in controllability.

[9]Pezzo and Pezzo's (2007) model distinguished between internal and external causes of an outcome with the idea that there is less motivation to defend against an externally caused negative outcome. The controllability measure in the present study (which was designed before the model was developed) may not have captured this idea fully.

DOI: 10.1002/acp

Of course, a number of differences exist between the present study and studies that support retroactive pessimism. We asked people to make and then recall predictions in a within-subjects design concerning a large number of real-world events that naturally differed in how controllable they were. Tykocinski and her colleagues instead asked participants in a lab setting to respond to a single hypothetical scenario in a between-subjects design in which she manipulated controllability. Further, some of our relatively low-controllable events (e.g. 'get very sick' 'be discriminated against'), though self-relevant and potentially threatening, may not produce the same sort of disappointment that Tykocinski and her colleagues had in mind (e.g. the loss of a sports team or political candidate). Any one of these factors might account for our different findings. Clearly, much remains to be learned about the complicated nature of motivational mechanisms affecting the hindsight bias.

Finally, there was a significant outcome controllability × happened interaction (Figure 3). For events that actually happened, those that were more controllable produced larger hindsight bias. This is consistent with past research showing that chance events are less susceptible to the hindsight bias (Wasserman et al., 1991). Because there are presumably fewer causal antecedents with uncontrollable outcomes, this finding provides indirect support for the idea that a successful sense-making process is necessary for hindsight bias to occur (Pezzo, 2003). As with valence, however, the effects of controllability only occurred for events that actually happened. As noted in Figure 3, events that did not happen still produced hindsight bias, but there was no difference in the size of this bias based on the event's controllability. Again, it is not clear to us how the sense-making process would operate on non-occurrences to produce such an effect. Nevertheless, a finding based on 3443 data points is rather compelling, and is worthy of future exploration.

Future research

Future studies of motivational mechanisms of hindsight bias should consider the possibility that motives might change *over time*. To date, only Renner (2003) appears to have examined such a possibility (although see Bryant & Brockway, 1997), arguing that motives to defend against a threatening outcome may diminish as participants 'come to peace' with that outcome. Although we took only one hindsight measure, we waited much longer (2–3 months) than do many laboratory studies, giving our participants considerably more time to make peace (or not) with self-relevant outcomes. Of course, the relative availability of explanations for the outcome that either reflect or incur blame will likely determine the extent to which any such motivational changes take place (Sanna, Schwarz, & Small, 2002). Unfortunately, in the present research, with such a large number of events, we were unable to solicit the explanations people considered for each of their outcomes.

Finally, researchers should consider the implications of exhibiting or not exhibiting hindsight bias. As one of us has recently argued (Pezzo & Pezzo, 2007), failure to show hindsight bias may not necessarily be a good thing. Indeed, it would seem to imply that one has not learned from that outcome. The present research suggests that, averaged across many events at least, people may be quite good at keeping themselves from accepting blame for their misfortunes. This may make them feel better, but it may not help very much in preventing future misfortunes.

DOI: 10.1002/acp

ACKNOWLEDGEMENTS

We thank Jan Peters, Hartmut Blank and Stephanie Pezzo for their helpful comments on an earlier version of the manuscript. Jill Rosa graciously assisted with collection and coding of the data for this study. This research was presented at the Hindsight Bias Workshop in Leipzig, Germany, July 8–10, 2005.

REFERENCES

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks: Sage.

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

Bryant, F. B., & Brockway, J. H. (1997). Hindsight bias in reaction to the verdict in the O.J. Simpson criminal trial. *Basic and Applied Social Psychology*, *19*, 225–241.

Brunswik, E. (1956). *Perception and a representative design of psychological experiments*. Berkeley, CA: University of California press.

Bryk, A. S., & Raudenbush, S. W. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed). Thousand Oaks, CA: Sage.

Burger, J. M., & Cooper, H. M. (1979). The desirability of control. *Motivation and Emotion, 3*, 381–393.

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, *48*, 306–307.

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

Darlington, R. B. (1990). *Regression and linear models*. New York: McGraw-Hill Publishing Company.

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.

Guilbault, R. L., Bryant, F. B., Brockway, J. H., & Posavac, E. J. (2004). A meta-analysis of research on the hindsight bias. *Basic and Applied Social Psychology*, *26*, 103–117.

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

Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, *107*, 311–327.

Hölzl, E., Kirchler, E., & Rodler, C. (2002). Hindsight bias in economic expectations: I knew all along what I want to hear. *Journal of Applied Psychology*, *87*, 437–443.

Louie, T. A. (1999). Decision makers' hindsight bias after receiving favorable and unfavourable 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.

Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Mark, M. M., Boburka, R. B., 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 foreseeability. *Memory*, *11*, 443–454.

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.

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

DOI: 10.1002/acp

Musch, J. (2003). Personality differences in hindsight bias. *Memory*, *11*, 473–489.
Ofir, C., & Mazursky, D. (1997). Does a surprising outcome reinforce or reverse the hindsight bias? *Organizational Behavior and Human Decision Processes*, *69*, 51–57.
Nezlek, J. B. (2001). Multilevel random coefficient analyses of event and interval contingent data in social and personality psychology research. *Personality and Social Psychology Bulletin, 27*, 771–785.
Pennington, D. C., Rutter, D. R., McKenna, K., & Morley, I. E. (1980). Estimating the outcome of a pregnancy test: Women's judgments in foresight and hindsight. *British Journal of Social and Clinical Psychology*, *19*, 317–324.
Pezzo, M. V. (2003). Surprise, defence, or making sense: What removes hindsight bias? *Memory*, *11*, 421–441.
Pezzo, M. V., & Beckstead, J. W. (2006). A multi-level analysis of rumor transmission: Effects of anxiety and belief in two field experiments. *Basic and Applied Social Psychology*, *28*, 91–100.
Pezzo, M. V., & Pezzo, S. P. (2007). Making sense of failure: A motivated model of hindsight bias. *Social Cognition*, *25*, 147–164.
Renner, B. (2003). Hindsight bias after receiving self-relevant health risk information. *Memory*, *11*, 455–472.
Roese, N. J., & Olson, J. M. (1996). Counterfactuals, causal attributions, and the hindsight bias: A conceptual integration. *Journal of Experimental Social Psychology*, *32*, 197–227.
Sanna, L. J., Schwarz, N., & Small, E. M. (2002). Accessibility experiences and the hindsight bias: I knew it all along versus it could never have happened. *Memory & Cognition*, *30*, 1288–1296.
Schkade, D. A., & Kilbourne, L. M. (1991). Expectation-outcome consistency and hindsight bias. *Organizational Behavior and Human Decision Processes*, *49*, 105–123.
Stahlberg, D., & Schwarz, S. (1999). Would i have known it all along if i would hate to know it? The hindsight bias in situations of high and low self esteem relevance. *Unpublished Research Report #99-34*, University of Mannheim.
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.
Tykocinski, O. E. (2001). I never had a chance: Using hindsight tactics to mitigate disappointments. *Personality and Social Psychology Bulletin*, *27*, 376–382.
Tykocinski, O. E., Pick, D., & Kedmi, D. (2002). Retroactive pessimism: A different kind of hindsight bias. *European Journal of Social Psychology*, *32*, 577–588.
Tykocinski, O. E., & Steinberg, N. (2005). Coping with disappointing outcomes: Retroactive pessimism and motivated inhibition of counterfactuals. *Journal of Experimental Social Psychology*, *41*, 551–558.
Verplanken, B., & Pieters, R. G. (1988). Individual differences in reverse hindsight bias: I never thought something like Chernobyl would happen. Did I? *Journal of Behavioral Decision Making*, *1*, 131–147.
Wasserman, D., Lempert, R. O., & Hastie, R. (1991). Hindsight and causality. *Personality and Social Psychology Bulletin*, *17*, 30–35.
Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin*, *25*, 1115–1125.
Wilson, T. D., Gilbert, D. T., & Centerbar, D. B. (2003). Making sense: the causes of emotional evanescence. In I. Brocas & J. Carrillo (Eds.), *The psychology of economic decisions. Vol. 1: Rationality and well being* (pp. 209–233). New York: Oxford University Press.
Winman, A., Juslin, P., & Bjorkman, M. (1998). The confidence-hindsight mirror effect in judgment: An accuracy-assessment model for the knew-it-all-along phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 415–431.

DOI: 10.1002/acp