Journal of Experimental Psychology: Applied
2012, Vol. 18, No. 2, e1–e16

1076-898X/12/$12.00 DOI: 10.1037/a0027959

Evolving Judgments of Terror Risks: Foresight, Hindsight, and Emotion: A Reanalysis

Baruch Fischhoff, Roxana M. Gonzalez, and Jennifer S. Lerner
Carnegie Mellon University

Deborah A. Small
University of Pennsylvania

The authors examined the evolution of cognitive and emotional responses to terror risks for a nationally representative sample of Americans between late 2001 and late 2002. Respondents' risk judgments changed in ways consistent with their reported personal experiences. However, they did not recognize these changes, producing hindsight bias in memories for their judgments. An intensive debiasing procedure failed to restore a foresightful perspective. A fear-inducing manipulation increased risk estimates, whereas an anger-inducing manipulation reduced them—both in predictions (as previously observed) and in memories and judgments of past risks. Thus, priming emotions shaped not only perceptions of an abstract future but also perceptions of a concrete past. These results suggest how psychological research can help to ensure an informed public.

Keywords: risk, emotion, terrorism, hindsight, availability

The events of September 11, 2001 confronted many people with a harsh reality that had previously been but a peripheral concern. It also initiated a learning process, first to acquire a basic understanding of the issues, then to follow them as they evolve over time (Fischhoff, 2002, in press). People need to learn how large the risks are, in order to balance protection from terrorism against other valued outcomes (e.g., economic cost, civil liberties, protection from other risks). They need to learn the processes creating and controlling terror, in order to think critically about proposals for managing it. They need to learn the "players," in order to determine whom to trust for information and advice. On each front, they need to update their beliefs and choices as experience accumulates.

Experimental psychology's research into judgment and learning should be relevant to understanding and aiding these choices. Indeed, there have been calls for its application (e.g., U.S. General Accounting Office, 2004). However, extrapolating to this arena poses challenges to the validity of laboratory research. Distinctive features of terror risks include high stakes, intense public discussion, and complex emotional content. A priori, each of these features might improve or degrade lay judgment, compared with the judgments observed in typical research settings. High stakes might increase people's desire to understand these risks and their feelings of accountability for doing so. However, high stakes can also degrade judgment, especially when people lack needed background knowledge, feedback, and cognitive skills (Camerer & Hogarth, 1999; Fischhoff, 1982; Gilovich, Griffin, & Kahneman, 2002; Lerner & Tetlock, 1999). Public discussion can provide useful information (Fischhoff, 2002; Slovic, 2001). However, its message is often ambiguous and uncertain, forcing citizens to distinguish facts from spin. Emotion can help people to integrate their beliefs and feelings (Gray, 2004). However, it can also leave them prey to transient affective states—and to manipulation by others (Loewenstein & Lerner, 2003).

In the present study, we examined three cognitive and affective processes in this distinctive context. We used a two-factor experimental design, embedded in the natural experiment created by September 11th and its aftermath. We sought generalizability by using a nationally representative sample of Americans. We sought ecological validity by having respondents participate in their homes to stimuli drawn from the national news media and presented through their televisions. We evaluated performance with both external standards (statistical estimates) and internal ones (within-person consistency over time).

The study is, in part, constructive replication, assessing the robustness of previously observed patterns in this context and, in part, extension, testing new hypotheses about the focal processes and their interactions. Three features of the experimental design make these extensions possible. First, the study simultaneously examined processes that are typically studied in isolation. Doing so reveals the relative strength of their effects and interactions in a design that has not been fine tuned to magnify a focal effect. Second, the study used both previously used measures, providing continuity with existing studies, and new ones, providing additional perspectives on the focal processes. Third, the juxtaposition raises new theoretical questions regarding (a) the effects of emotions on judgments of the past, (b) the efficacy of an ambitious debiasing procedure, and (c) the roles of different experiences on belief updating.

Baruch Fischhoff, Roxana M. Gonzalez, Jennifer S. Lerner, Department of Social and Decision Sciences, Carnegie Mellon University; Deborah A. Small, Wharton School, University of Pennsylvania.

This research was supported by National Science Foundation Grants SES-0201525, PECASE SES-0239637, and SBR-9521914; National Institute of Mental Health Grant MH62376; and American Psychological Association (Division 9) grants. We thank Wändi Bruine de Bruin, Robyn Dawes, Mike Denis, Julie Downs, James Gross, Rick Li, and Bill McCready for their help. The views expressed are those of the authors.

Correspondence concerning this article should be addressed to Baruch Fischhoff, Department of Social and Decision Sciences, Carnegie Mellon University, Porter Hall 208, Pittsburgh, PA 15213. E-mail: baruch@cmu.edu

Addressing these multiple goals inevitably creates complex designs and data. To simplify the exposition, we first describe the three focal processes in general terms, then we consider the theoretical issues studied here and the applied implications of the anticipated effects.

Theoretical Background

Availability-Based Judgment

It is natural to judge an event's likelihood by how easily examples come to mind (Tversky & Kahneman, 1973). Such *availability* can be a valid cue. Events are often more likely in the future if they have frequently been experienced or observed in the past. However, availability can also mislead when events are disproportionately visible and observers cannot adjust for that sampling bias. For example, the news media report homicides at a higher rate than suicides, which could account for people overestimating homicides relative to suicides (Combs & Slovic, 1979; Lichtenstein, Slovic, Fischhoff, Layman, & Combs, 1978). Predicting availability-based judgment, as well as distinguishing it from other memory-based processes, requires specifying, a priori, how the heuristic is used (e.g., Beyth-Marom & Fischhoff, 1977; Schwarz & Vaughan, 2002; Viscusi & O'Connor, 1984). Assessing the accuracy of such judgments requires having comparable estimates of actual occurrence rates.

In one early, and often-cited, study assessing the role of availability in frequency judgments for real-world events, Lichtenstein et al. (1978) asked respondents to estimate the annual U.S. death toll from 41 causes of death, offering either annual deaths from motor vehicle accidents (MVA) or electrocution (E) as an anchor. A second group of respondents reported the availability of these events, operationalized as their closest experience with such deaths to a close friend or relative, to someone else they knew, and to someone they had heard about. Mean scores on this 1–3 scale were correlated with log geometric mean death estimates (−.90, MVA; −.88, E) and with residuals from the best-fit line predicting those judgments from log statistical estimates (−.19, MVA; −.19, E). Respondents' reported experiences with these events as sources of suffering had weaker correlations with the death estimates (−.52, MVA; −.50, E) and similar correlations with the residuals (−.22, MVA; −.16, E). Thus, availability, as defined by self-reported experience, predicted frequency judgments both overall and in terms of deviations from frequency judgments for events with similar statistical estimates (i.e., the residuals).

Hindsight-Based Judgment

It is also natural to integrate new observations with existing knowledge to create a comprehensive mental model (Bartlett, 1932). That integration process creates a new *gist* (or *semantic*) memory of the event, while leaving a *verbatim* (or *anecdotal*) memory, typically fading with time (e.g., Reyna, 2004; Reyna & Brainerd, 1991, 1995). Not recognizing such changes creates hindsight bias: When people try to recall past beliefs, their view is unwittingly contaminated by subsequently acquired ones. As a result, they exaggerate the extent to which they knew all along what they actually learned only later (Fischhoff, 1975; for related accounts, see Hoffrage, Hertwig, & Gigerenzer, 2000; Wilson & Brekke, 1994). One concern in hindsight bias studies is having respondents misrepresent their prior beliefs, in order to exaggerate what they once knew. A task designed to reduce that risk challenges respondents to recall their own prior predictions (Fischhoff & Beyth, 1975). Few studies have used this task with long periods between prediction and recall, presumably because of the logistical difficulty of finding respondents the second time. The longer the time period, the greater should be reliance on gist memory, enhancing its natural preferential status (Reyna, 2004).

Christensen-Szalanski and Wilham's (1991) meta-analysis found an average bias of approximately 10% difference between actual predictions (in foresight) and recalled ones (in hindsight). The bias was larger for events believed to have occurred, perhaps because their greater salience evokes deeper processing, compared with nonevents, which might not even be noticed. Guilbault, Bryant, Brockway, and Posavac's (2004) meta-analysis found a mean effect size of $d = 0.39$ (95% confidence interval [CI] = .36–.42), with smaller effects for studies using real-world events and nonneutral ones, but no difference with various debiasing manipulations.

Establishing that a difference between original and recalled judgments represents hindsight bias requires measuring belief change. One possible measure is the use of the same observations that inform availability-based judgments (Fischhoff & Beyth, 1975). Other things being equal, observing an event should increase its retrospective, or *postdicted*, likelihood (Fischhoff, 1975). However, related beliefs regarding the general situation may also change. As a result, the postdicted probability of an event that occurred could be even lower than the original one, depending on what else has been learned. ("Originally, I thought that the event was fairly likely. Its happening, initially, made it seem even more likely. However, now that I understand things more fully, I realize that it was not as likely as I had thought.") Such changes need to be assessed with questions such as, "Knowing everything that you know now, what was the probability that the event would happen?" In principle, people could feel that an event is truly surprising, eliminating (or even reversing) the bias, although demonstrations are sparse and controversial (e.g., Hoch & Loewenstein, 1989; Mark & Mellor, 1994; Mazursky & Ofir, 1990, 1996).

Emotion-Based Judgment

Emotions provide potentially useful inputs to judgment and choice. Emotions can serve as *somatic markers*, providing knowledge about events that is not readily verbalized (Damasio, 1995). People may generate sympathetic nervous system responses to a risky option before they can articulate its risk (Bechara, Damasio, Tranel, & Damasio, 1997). As a result, emotions might help people to gauge future risks and recall past ones—hypotheses we examined in this study. However, emotions can also create transient shifts in perception, which may influence choice in unnoticed or unwanted ways (Loewenstein & Lerner, 2003; Loewenstein, Weber, Hsee, & Welch, 2001; Mellers, Schwartz, & Ritov, 1999; Schwarz & Clore, 1996; Slovic, Finucane, Peters, & MacGregor, 2002). For example, incidental anger from one situation can increase optimism in others, perhaps mobilizing needed resources or perhaps creating unwarranted and unrecognized feelings of invulnerability (Lerner & Gonzalez, 2005; Lerner, Gonzalez, Small, & Fischhoff, 2003; Lerner & Keltner, 2000, 2001). These effects of fear and anger can be explained within the *appraisal tendency* framework (Lerner & Keltner, 2000), in which each emotion activates a specific cognitive predisposition when people appraise future events. These appraisal tendencies are goal-directed pro-

cesses by which emotions affect judgment and choice in ways specific to the events that evoke them.

In a study described more fully below, Lerner et al. (2003) found that respondents exposed to a fear-inducing manipulation assigned, on average, a 7.8% ($d = 0.26$) higher probability to five negative consequences of terror than did respondents exposed to an anger-inducing manipulation. These emotions carried over to probability judgments for routine risks having no obvious connection to the terrorism-related manipulations (e.g., getting the flu). The mean difference here was 6.8% ($d = 0.30$).

Implications for Responding to Terror

Thus, these processes can be adaptive but they can also bias judgments. Availability can misdirect priorities by unduly highlighting some risks. For example, critics of the Iraq war argue that focusing on Iraq exaggerated those risks while diverting resources from the larger threat of al Qaeda and its allies. Hindsight bias can make people unduly critical of themselves and their leaders—feeling remorse and blame while underestimating how little was known and how much there is to learn. Anger's unrecognized effects can exaggerate the perceived effectiveness of personal and public actions (e.g., going to war); fear can do the opposite. Such emotions can occur naturally or be induced by those hoping to manipulate public concerns (Rich, 2004). If emotions change over time, then that element will be missing when people try to retrieve past perspectives, increasing hindsight bias and leaving them wondering why they made past choices (e.g., how they felt when deciding whether to support the Iraq war).

The size of these effects is, naturally, context dependent as is their impact (von Winterfeldt & Edwards, 1986). For example, an average hindsight bias of 10% might be large enough to shift some choices from appearing sound to appearing ill advised (e.g., originally thinking that the chance of weapons of mass destruction in Iraq was 11% vs. recalling it as having been 1%). Other decisions may be insensitive to the same change (e.g., if a 30% chance of disruption was enough to make one avoid flying in late September 2001, a 20% recalled chance might have led to the same choice). A 10% emotion-induced change in risk judgments might tip some close decisions and be irrelevant to clear-cut ones.

In addition to their decision-specific impacts, such effects have broader implications for societal decision making. Citizens' role in managing a risk should depend on how well they understand it—and on how well they could understand it given diligent efforts to inform them. Underestimating citizens' competence undermines democratic processes by wrongly casting doubt on their ability to participate. Conversely, overestimating citizens' competence can mean demanding unrealistic cognitive mastery (Fischhoff, 1995, 1999, 2002; Leiss & Chociolko, 1997; Lupia, 1994; National Research Council, 1989; Slovic, 2001), denying them needed help. The sensitivity of citizens' beliefs to emotional manipulations shows how vulnerable democratic processes are to that threat. This study sheds light on this general question, too.

The Study

Overview

In late 2001, Lerner et al. (2003) studied a nationally representative sample of 973 Americans (see also Fischhoff, Gonzalez, Small, & Lerner, 2003b). Here, we compare judgments made at that time with judgments made by a subset of the same individuals a year later. The first wave of the study was conducted in November–December 2001 during the anthrax crisis and the Afghanistan campaign. The second wave was conducted in November–December 2002, as the abortive U.S. smallpox vaccination campaign was emerging, Iraq was being inspected, and the Mombasa and Bali attacks on tourists were fresh news (Fischhoff, Bruine de Bruin, Perrin, & Downs, 2004). This context allows us to test the following hypotheses regarding the robustness of the three focal processes and their relationships.

Availability-Based Judgment

In November 2001, terrorism was relatively unfamiliar to most Americans, including many officials. Given the intense public interest in terror, the ensuing year could have prompted extensive belief change, reflecting direct observation, media reports, study, and conversation. To measure that learning, we elicited respondents' experience with terrorism for several events (e.g., injury, anthrax) and several degrees of immediacy (e.g., their own lives, close friends, relatives). If respondents rely on availability (as expressed in these measures) when updating their beliefs, then their risk judgments should correlate with their reports. Of course, such correlations cannot prove reliance on availability. For example, personal experience with a risk might lead people to seek relevant information about it. The additional information could then make the risk seem more likely, perhaps amplified by mere exposure effects.

We took these measures of availability from Lichtenstein et al. (1978) and anticipated a constructive replication of their results. That is, we predicted that risk judgments would be correlated with the availability of observations, with the strongest relationships for the most direct observations. Unlike Lichtenstein et al., we elicited both risk judgments and availability reports from the same respondents, allowing within-event tests of these hypotheses for each measure of availability. These correlations resemble those that Lichtenstein et al. observed between availability judgments and the residuals from the best-fit line predicting judgments from statistical estimates (representing differences in risk judgments for events with similar statistical frequency). Those correlations ranged from .16 to .22, although the events in the two studies are so different as to preclude anything but the weakest inferences about effect size.

The longitudinal design also allows evaluating the accuracy of respondents' judgments in late 2001 regarding their predicted experiences over the ensuing year. As mentioned, Lichtenstein et al. (1978) found correlations of around .90 between log (geometric mean) frequency judgments and log statistical estimates (where the latter ranged over almost six orders of magnitude). Although people typically show some insight into risk levels (Fischhoff, Bostrom, & Quadrel, 2002), it would be noteworthy if that were the case with such novel events where people lack consensual theories for the processes creating and controlling them (Fischhoff et al., 2004).

Hindsight-Based Judgment

If beliefs have changed, as the result of experience, then the resulting (hindsight) state of belief poses a potential barrier to reconstructing previous (foresight) states of belief. Unless they overcome that barrier, people might not understand why they made personal choices (e.g., traveling less) or endorsed public policies (e.g., the U.S. Patriot Act). As a measure of respondents' ability to

access their foresight beliefs, we asked them to recall their original judgments, emphasizing memory accuracy (following Fischhoff & Beyth, 1975).

We assessed the extent of bias in two ways, reflecting different ways of measuring belief change. One (also following Fischhoff & Beyth, 1975) uses self-reported personal observation; bias occurs when respondents remember having assigned higher probabilities than they actually did to events that had happened and lower probabilities to events that had not. We represented those observations by the availability measure that showed the strongest correlations with the probability judgments. (As mentioned, we expected that measure to be personal observation.)

The second way of assessing hindsight bias used a new measure: respondents' 2002 *postdictions* for the events that they had predicted in 2001. Specifically, we asked respondents what they now believed the probability of each event to have been, taking into account all that they now knew. If respondents saw the world as having been safer than they originally had (postdiction $<$ original prediction), then hindsight bias means recalling having seen less risk than they had predicted (memory $<$ original prediction). If respondents saw the risk as having been higher than it had originally seemed, memories should be biased the other way. Because postdictions can accommodate all that respondents have learned, they provide a more comprehensive measure of belief change than respondents' event-specific observations. However, because postdiction might be seen as an intellectually challenging task, we examined these judgments' construct validity to help readers assess the contribution of this new measure.

A priori, we saw reasons for predicting that the circumstances around terror would both decrease hindsight bias (e.g., if media coverage of the evolving situation increased awareness of learning) and increase it (e.g., if experienced-based belief change prompted deep processing of observations). As a result, we predicted effects like those generally seen—an average difference of 10% between predicted and recalled probabilities with greater bias for events than for nonevents. We realized that beliefs about specific risks would be embedded in beliefs about the overall situation in which the risk level might seem to change over time. However, without knowing what those general changes were, we had no prediction for their effects.

Emotion-Based Judgment

Terror can evoke multiple, conflicting emotions, most notably fear and anger. Appraisal-tendency theory predicts that evoking anger will increase feelings of certainty and individual control whereas evoking fear will enhance feelings of uncertainty and situational control (Lerner & Keltner, 2001; Smith & Ellsworth, 1985). The 2001 study supported these predictions, in the judgments of respondents randomly assigned to conditions that primed either fear or anger, respondents in the anger condition saw less risk from terrorism than did those in the fear condition, even though, of course, they shared the same world. Risk judgments were similarly correlated with naturally occurring emotions, as expressed in a separate survey conducted shortly after September 11th (and about two months before the first wave of our study). A third group of respondents received a sadness manipulation, included for other theoretical reasons. Lerner et al. (2003) did not analyze its effects, nor will we here.

Late in 2002, the second wave of the study randomly assigned half of the respondents to repeat their 2001 emotion induction (the relive condition), while the other half (neutral) completed a task intended to maintain emotional neutrality. All respondents then proceeded to the judgment tasks. This design allowed us to make four predictions about affective processes:

1. People will be as susceptible to the effects of the emotion manipulations in 2002 as they were in 2001. Appraisal-tendency theory predicts higher risk estimates in the fear condition than in the anger condition. However, we had conflicting hypotheses regarding the size of those effects relative to those in 2001. One hypothesis is that being further from the September 11th attacks will reduce ambient cues, thereby increasing the manipulation's impact (Forgas, 1995). If so, then the manipulations should have stronger effects on judgment in 2002 than they had in 2001. The contrary hypothesis is that the passage of time will have reduced concern over terrorism, thereby reducing respondents' susceptibility to the terror-related stimuli used in the induction. If so, then the effects should be weaker than in 2001. In the absence of previous studies manipulating distance from a focal event, we predicted that these factors would cancel out one another, leaving emotion effects equally potent a year later.

2. The emotion manipulation will reduce hindsight bias by restoring respondents' prior state of mind before they recall their predictions. Several lines of research support this hypothesis. Studies of hot/cold empathy gaps find that people can reconstruct their past decisions better if they can recreate their prior emotional state (Loewenstein, 1996; Loewenstein & Adler, 1995; Van Boven, Dunning, & Loewenstein, 2000; Wilson & Gilbert, 2003). Studies of somatic markers find that emotions carry implicit signals shaping judgments and decisions (Bechara et al., 1997; Damasio, 1995). Reliving emotions should retrieve those implicit signals and, with them, the ability to recall the earlier judgments. Reactivating emotions has long been known to reactivate previously associated memories (for an early review, see Bower, 1982). By priming respondents' original emotions, our manipulations should also prime the associated memories, thereby recreating the original fear/anger difference. If the emotion manipulations also reinstate beliefs that were held a year earlier, then they should reduce hindsight bias. Neutral condition respondents should judge the risks similarly, whichever emotion induction they had received a year earlier, barring a long-term carry-over effect. Davies (1987) had some success in reducing hindsight bias by having respondents review written records of their previous thoughts without deliberately considering emotional content. As a result, we predicted reduced hindsight bias with a manipulation that also evoked past emotions.

3. The emotion inductions will color judgments of past risks as they have judgments of future risks. In other contexts, cues provided by actual experience have substantially reduced emotion carry-over effects (Forgas, 1995). As a result, we predicted that judgments of the past would be less sensitive to incidental emotion effects, but sensitive nonetheless.

4. The fear and the anger inductions will have equal effects on risk judgments, relative to a neutral baseline. The original study had no neutral condition, leaving these relative impacts unclear. (We thought it impossible to create a neutral condition at that time, given the apparent intense mix of fear and anger in American society two months after the attacks.) The fear and anger manipulations had, however, similar effects on self-reported emotions. As a result, we expected the two manipulations to affect risk

judgments equally relative to those reported in the neutral condition. A recent study (Lerner & Gonzalez, 2004) found that, compared with a neutral baseline, fearful individuals made pessimistic judgments on various life domains, whereas angry individuals made optimistic judgments (e.g., career, relationships, social competence).

Summary

We examine the effects of experience, memory, and emotion on judgment in the intense, novel context of terrorism, as these processes evolved from 2 months after the September 11th attacks until a year later. We use a two-factor experimental and longitudinal design with a random sample of Americans responding in their homes to stimuli drawn from life. The design allowed us to test the robustness, interactions, and relative strength of processes previously observed in isolation with more typical experimental settings and samples. We used a set of interrelated old and new measures providing both replication and extension. We validated two measures independently of their use in hypothesis testing (the measure of experience used to assess hindsight bias, the effects of the emotion manipulation used to reduce hindsight bias).

Method

Design

In November 2001, respondents were randomly assigned to one of two emotion conditions: fear or anger. The manipulation involved having respondents write a short passage describing aspects of the September 11th attacks that evoked the target emotion for them in a way that would also evoke it for other people, followed by viewing a picture and hearing an audio clip drawn from the national news media that had been found, in pretests, to prime that emotion. The judgment tasks came next. (For additional details, see Lerner et al., 2003.)

One year later, respondents were randomly assigned to one of two recall conditions: relive or neutral. Relive respondents repeated their 2001 emotion manipulation before judging risks. After being instructed to experience fully the emotions that they had felt in 2001, they reread the passages that they had written, viewed the same picture, and heard the same audio clip. To equate time on task, we asked neutral respondents to review their materials from 2001 (i.e., the written response, picture, and audio clip), while adopting a neutral, objective, analytical perspective; they then performed the judgment tasks. Note that the neutral condition itself includes a debiasing manipulation, insofar as it helps respondents to recall what they were thinking when they answered the judgment tasks a year earlier (cf. Davies, 1987). The relive condition adds the attempt to reinstate emotions from that time.

Dependent Measures

Risk judgments. In 2001, respondents judged the probability that each of five terror-related risks and three routine risks (see Table 1) would occur in the next 12 months using an open-ended scale anchored at 0% (*the event is impossible*) and 100% (*the event is certain to happen*). We labeled these judgments *Prediction 2001*. Probability judgments for screening mail and getting the flu had normal distributions. For the other six events, risk judgments were positively skewed (defined as having a skewedness statistic twice the standard error), with skewedness statistics ranging from 1.03 to 1.37 (SE = .08). As a result, we used nonparametric statistical tests when considering individual events. Averaging across events produced a more normally shaped distribution (skewedness statistic = .09, SE = .08).

In 2002, respondents made three judgments for each event in the following order, completing all eight events before proceeding to the next judgment:

Recall 2001:	"Please remember what you predicted a year ago or, if you cannot remember, write what you would have said then."
Update 2001 (postdiction):	"Please estimate what the probability was, knowing what you know now about the United States and its enemies."
Prediction 2002:	The same as Prediction 2001, but for the ensuing 12 months.

This order follows the natural chronology of the three perspectives. With it, no response precedes Recall 2001, thereby providing the clearest assessment of hindsight bias.

Experience.

Table 1
Probability Judgments for Risk Events (In Order of Task Completion)

Event	Prediction 2001			Recall 2001			Update 2001			Prediction 2002			% reporting personal experience
	M	*Mdn*	*SD*	*M*	*Mdn*	*SD*	*M*	*Mdn*	*SD*	*M*	*Mdn*	*SD*	
Terror risks													
Being hurt	20.3	10	22.3	16.9	5	23.0	14.6	5	21.6	19.2	10	25.8	0.3
Traveling less	33.6	15.6	37.0	20.7	4.3	30.6	23.4	5	33.0	26.9	10	33.5	30.0
Trouble sleeping	23.5	10	30.0	11.4	1	19.9	12.9	2	22.3	14.6	5	22.5	22.0
Screening mail	55.3	60	39.4	19.5	5	29.9	23.2	5	32.1	29.9	10	34.6	34.0
Anthrax antibiotics	22.4	5	30.5	10.7	1	20.0	9.6	0	18.3	14.5	5	21.9	0.7
Routine risks													
Getting flu	46.9	50	31.8	30.9	20	28.8	30.2	20	31.1	34.5	25	29.2	28.0
Other violent crime	22.2	10	23.3	17.9	10	21.9	14.4	5	19.6	15.8	10	18.9	0.9
Dying	33.0	20	33.4	22.4	10	24.5	19.1	5	24.1	21.6	10	23.4	0.0

Note. Table 1 pools the control and relive conditions. Looking separately at the different groups, our conclusions would not be different for either recall or emotion conditions. Prediction 2001 presents predictions made in November 2001 for the next 12 months. Recall 2001 presents respondents' memories, in 2002, of their 2001 predictions. Update 2001 presents respondents' best guesses at what each risk really was for the preceding 12 months. Prediction 2002 presents predictions made in November 2002 for the next 12 months.

After the probability judgments, respondents indicated whether, over the past year, each risk event had occurred (a) to them personally, (b) to a close friend or relative, (c) to someone else they knew, and (d) to someone they had heard about. For the event of "dying," we omitted the self and the heard-about reports. The wording followed that used by Lichtenstein et al. (1978).

Emotion self-reports.

At the end of the experiment, respondents completed 5-item scales, reporting how much they felt each of the two focal emotions during the task preceding the judgment tasks. Response options ranged from 0 (*do not feel the emotion the slightest bit*) to 8 (*feel the emotion even more strongly than ever before*). Responses indicated coherent scales, with some respondents reporting more of each emotion across the five relevant questions (2001, $\alpha = .94$, fear, $\alpha = .94$, anger; 2002, $\alpha = .95$, fear, $\alpha = .95$, anger). For subsequent analyses, we used average responses on each scale.

Sample

Source. Respondents belonged to a 75,000 household nationally representative panel maintained by Knowledge Networks (Menlo Park, CA). The panel tracks U.S. Census statistics on key demographic variables, such as age, race, ethnicity, geographical region, employment status, income, and education (Krotki & Dennis, 2001). Panel households receive free WebTV and Internet access in return for completing 10–15 min Internet surveys 3–4 times per month. With other tasks, panel members have been found to respond similarly to first-time respondents (Denis, 2001).

Recruitment.

Potential respondents received notice of our study through a password-protected e-mail account, with 2 weeks to provide their responses. Prior to beginning the study, respondents (or their parents, for those under age 18) provided informed consent. Responses are confidential, with identifying information never revealed without respondent approval. Respondents can complete a study only once and may stop at any time without affecting their WebTV or Internet service. Respondents have individual passwords so that the same person should be responding both times and parental intervention is needed before minors can continue. For further details, see http://www.knowledgenetworks.com/ganp/.

2001 study.

Knowledge Networks contacted 1,786 panel members, inviting participation from people who could spend 20 uninterrupted minutes; 1,030 individuals opened the study between November 10 and November 29, 2001. The study sample included the 973 who answered almost all questions. Their demographics roughly matched U.S. Census figures, allowing results to be generalized to the U.S. population.

2002 study.

Between November 15 and December 30, 2002, Knowledge Networks sent notice of the new study to the 869 members of the 2001 sample still in the panel. Among them, 81.1% acknowledged receiving the invitation; 82.6% of those who acknowledged receipt consented to participate, for an overall response rate of 67.0%. The final sample had 532 adults (259 men and 273 women) and 50 teens (24 men and 26 women). An attempt was made to recruit 2001 respondents who had dropped out of the Knowledge Networks panel. However, the response rate was so low that the few individuals recruited in this way were not included in this data set.

The only statistically significant demographic difference between the 2001 and 2002 samples was age. In part because respondents were a year older (in 2002), and in part because rerecruitment was a bit more successful with older respondents (perhaps because they had more stable residences). Despite the demographic comparability of the 2001 and 2002 samples, it is still possible that those who decided to participate differ in some way from the overall sample insofar as the e-mail message regarding the 2002 study reminded them about their participation in the 2001 survey in addition to describing the topic of the 2002 study.

For each time period (2001, 2002), results are adjusted statistically to represent the general population. Specifically, in order to correct for any nonresponse bias, representative samples were selected by poststratification weighting of the panel to match benchmarks from the most recent U.S. government statistics for sex, age, race, ethnicity, education, and region. Samples are drawn with probabilities proportional to the panel weights, using a systematic sample applied to eligible panel members. Eligible panel members resemble the national population distributions for key demographic variables within sampling error. Across time periods, respondents serve as their own controls.

Results

Overview

After introducing the risk judgments, we analyze, in turn, availability, hindsight, and affective processes, and the interactions between them. We conclude with the debiasing effects (if any) of reliving the emotion induction.

Risk Judgments

Table 1 pools responses for the neutral and relive conditions, after analyses of variance revealed no significant Emotion $\times$ Recall Factor interactions, for any of the three judgment tasks: Recall 2001 items, $F(1, 371) = 2.87$, *ns*, $MSE = 279.50$, Cohen's $f = .05$; update items, $F(1, 371) = 2.13$, *ns*, $MSE = 264.48$, Cohen's $f = .05$; prediction items, $F(1, 371) = 0.19$, *ns*, $MSE = 284.71$, Cohen's $f = .05$. Given the skewed distributions (range of skewedness statistic for Prediction 2001 items = 0.80–1.93, $SE = .11$; Recall 2001 items = 0.76–2.11, $SE = .10$; Update 2001 items = 0.92–2.26, $SE = .10$; Prediction 2002 items = 0.87–2.56, $SE = .10$), we present medians, as well as means.

Table 1's first columns repeat the Prediction 2001 judgments from Lerner et al. (2003). The last column shows the percentage of respondents who reported (in 2002) having had each experience during the preceding year. Most events occurred less often than the expectation implied by the mean (or median) prediction probability. The two exceptions (traveling less, having trouble sleeping) involved events that some respondents might already have been experiencing when making their 2001 judgments. Thus, if these probability judgments and self-reports are taken literally, respondents experienced a much safer world than they had anticipated.

The Update 2001 columns present respondents' postdictions in late 2002 of what the preceding year's risks had been, given their current beliefs about the period. Each risk was seen as having been significantly less likely than it had seemed before the year had begun (Update 2001 $<$ Prediction 2001; see Table 2 for statistical tests). That change in respondents' beliefs is consistent with their reported experience (of little happening).

As a check on respondents' understanding of the postdiction task, we looked at whether they viewed their experiences as having been inevitable. When respondents reported an event, they said that it had been certain to happen (Update 2001 = 100%) 21.1% of the time (CI: 19.8%, 22.4%). When an event was not reported,

Table 2
Statistical Comparisons of Prediction 2001 With Recall 2001, Update 2001, and Prediction 2002 Judgments for Risk Events and Statistical Comparisons of Update 2001 With Prediction 2002 Judgments for Risk Events

	Comparisons with Prediction 2001									Comparison with Update 2001		
	Recall 2001			Update 2001			Prediction 2002			Prediction 2002		
Risks and event	*df*	*Z*	*d*	*df*	*Z*	*d*	*df*	*Z*	*d*	*df*	*Z*	*d*
Terror risks												
Being hurt in a terror attack	564	−4.21**	0.15	561	−6.35**	0.26	562	−2.46*	0.05	589	−6.66*	0.19
Traveling less	562	−8.07**	0.38	549	−7.39**	0.29	562	−4.62**	0.19	580	−5.46**	0.11
Trouble sleeping	565	−10.64**	0.48	555	−9.53**	0.40	563	−6.75**	0.34	574	−4.15**	0.08
Screening my mail	559	−16.39**	1.02	550	−15.31**	0.89	555	−13.38**	0.69	574	−8.16**	0.20
Taking antibiotics against anthrax	563	−8.96**	0.45	553	−9.98**	0.51	560	−5.38**	0.30	576	−9.17**	0.24
Routine risks												
Getting the flu	565	−10.31**	0.53	554	−9.91**	0.53	561	−8.24**	0.41	578	−5.93**	0.14
Being a victim of violent crime (other than terror)	561	−4.46**	0.19	552	−7.56**	0.36	558	−6.12**	0.30	578	−4.58**	0.07
Dying	553	−6.14**	0.36	542	−8.01**	0.48	550	−6.52**	0.40	574	−5.68**	0.11

Note. Asterisks compare Prediction 2001 with the three other sets of judgments (Recall 2001, Update 2001, and Prediction 2002), or they compare Update 2001 with Prediction 2002 (using Wilcoxon matched-pairs signed-rank tests, given the skewed distributions).
* $p < .05$. ** $p < .001$.

respondents thought that it had still been possible (Update 2001 ≠ 0%) 54.4% of the time (CI: 52.8%, 55.9%). Thus, respondents were not determinists; rather, they recognized that there had been uncertainty about what would happen even after knowing how things turned out. Thus, the difference between Prediction 2001 and Update 2001 judgments arguably captures some of what respondents had learned about the risks that they had faced beyond what they learned from their personal experiences.

The Prediction 2002 results shows risk judgments for the next 12 months. For each event, respondents saw significantly lower risks in 2002 than they had seen in 2001 (Prediction 2002 < Prediction 2001; see Table 2 for tests). That change is consistent with their collective personal experience of a safer-than-expected world. Perhaps tellingly, respondents saw a relatively similar (although still statistically lower at $p < .05$) chance of being hurt by terror in the 2nd year, despite having had almost no such personal experiences in the 1st year. Consistent with this constant general fear, respondents saw the next year as posing significantly greater risk than the previous one had (Prediction 2002 > Update 2001, $p < .001$, for each risk; see Table 2).

At this aggregate level, these belief changes were consistent with respondents' self-reported experiences. The next section asks whether this is also true at the individual level and, if so, which measure predicts risk judgments most strongly: direct personal experiences or observations of others' fate (which are less directly relevant but based on larger samples).

Availability Measures as Predictors of Risk Judgments

Table 3's first column shows respondents' reported experiences arranged by decreasing proximity. (The results for self also appeared in Table 1.) Generally speaking, there were more reports of the less direct experiences fitting the larger number of people to whom each could have happened. The exceptions were events where people might not reveal their personal experiences to others (e.g., having trouble sleeping).

The remaining columns show correlations between respondents' probability judgments and whether they reported each experience (0,1; using Kendall's Tau_b for the statistical tests, given the skewed probability distributions).

For Prediction 2001, positive correlations mean that respondents who saw risks as more likely in November 2001 were also more likely to report such experiences in the following year. All significant correlations were positive, indicating that respondents had some insight into their relative risk. In some cases (e.g., trouble sleeping, traveling less), they might have already been experiencing the event. In some cases (e.g., trouble sleeping, traveling less), a broader definition of the event may have justified a higher probability judgment. Generally speaking, more direct exposures had higher correlations, meaning that respondents' judgments predicted their own experiences better than those of others. Rare events (e.g., being hurt) leave little to predict.

For Update 2001, positive correlations mean that respondents who reported an event also saw themselves as having faced a greater risk of the event happening. Such correlations were found for most events (except crime and dying) consistent with respondents inferring their risk from their experience. The strongest correlations were with direct personal experience. Indeed, adding observations of others' experiences did not significantly increase the correlations with risk judgments.

For Prediction 2002, positive correlations mean that respondents who reported an event also saw greater risk in the future. Here, too, experience predicts risk judgments, with the strongest correlations for direct personal experience. Considering indirect experiences, again, does not significantly increase the correlation with risk judgments.

Summary

Respondents' predictions in 2001 correlated with their subsequently reported personal experiences, showing some apparent insight into their relative degree of risk. In cases where an event's definition might be ambiguous (e.g., trouble sleeping), respondents are evaluated by their own standard. These personal experiences also correlated with respondents' judg-

Table 3
Respondents' Reported Experience With Risk Events and Correlation With Probability Judgments

Event	% reporting	Prediction 2001		Update 2001		Prediction 2002	
		Tau_b	R^2	Tau_b	R^2	Tau_b	R^2
Being hurt in a terror attack							
Self	0.3	−.04	.00	.03	.00	−.02	.00
Close friend or relative	3	.19**	.09	.21**	.11	.07*	.01
Other than close friend or relative	8	.08	.02	.16**	.06	.04	.00
Heard of anyone	57	−.01	.00	−.03	.00	−.09	.02
Traveling less							
Self	30	.29**	.19	.49**	.48	.47**	.45
Close friend or relative	33	.21**	.11	.41**	.36	.41**	.36
Other than close friend or relative	31	.17**	.07	.26**	.16	.23**	.13
Heard of anyone	50	.17**	.07	.23**	.12	.20**	.10
Trouble sleeping							
Self	22	.23**	.13	.41**	.36	.38**	.32
Close friend or relative	28	.20**	.10	.27**	.17	.26**	.16
Other than close friend or relative	19	.13**	.04	.20**	.10	.17**	.07
Heard of anyone	33	.12*	.04	.16**	.06	.15**	.05
Screening my mail							
Self	34	.31**	.22	.55**	.58	.45**	.42
Close friend or relative	33	.24**	.14	.43**	.39	.37**	.30
Other than close friend or relative	25	.13**	.04	.25**	.15	.18**	.08
Heard of anyone	43	.11*	.03	.18**	.08	.13**	.04
Taking antibiotics against anthrax							
Self	0.7	.12*	.04	.31**	.22	−.04	.00
Close friend or relative	2	−.05	.01	−.06	.01	−.07	.01
Other than close friend or relative	2	−.05	.01	−.04	.00	−.05	.01
Heard of anyone	21	−.03	.00	−.06	.01	−.08	.02
Getting the flu							
Self	28	.33**	.25	.41**	.36	.42**	.38
Close friend or relative	63	.22**	.12	.25**	.15	.28**	.18
Other than close friend or relative	59	.34**	.26	.22**	.12	.17**	.07
Heard of anyone	73	.05	.01	.12**	.04	.07	.01
Being a victim of violent crime (other than terror)							
Self	0.9	−.02	.00	−.01	.00	−.03	.00
Close friend or relative	6	−.03	.00	−.01	.00	−.02	.00
Other than close friend or relative	13	−.02	.00	−.02	.00	.03	.00
Heard of anyone	51	−.07	.01	−.03	.00	−.02	.00
Dying							
Close friend or relative	46	−.02	.00	.11*	.03	−.07	.01
Other than close friend or relative	60	−.03	.00	.02	.00	−.01	.00

Note. Asterisks represent significant correlations between whether respondents reported each exposure (0,1) and their probability judgments (using Kendall's Tau_b, given the skewed distributions). We followed the procedures outlined by Walker (2003) in assessing the effect size (using R^2) for the Kendall's Tau_b correlations.
* $p < .05$. ** $p < .001$.

ments of what their risk had been (Update 2001) and would be for the following year (Prediction 2002). Considering respondents' observations of others' experiences did not improve predictions of their risk judgments, even though those observations reflected larger samples of experience. Thus, with these real-world events and this representative sample, availability—defined as self-reported personal experience—both predicts future risk judgments and is predicted by past ones. Although availability can be confounded with other cognitive and affective processes, this simple measure provides a parsimonious account.

Hindsight Bias

The world proved safer than respondents had predicted late in 2001. It seemed safer as well when they made new predictions late in 2002. Such belief change creates the possibility of hindsight bias, unless respondents can recall how risky the world once looked. Comparing the first and second columns of Table 1 shows hindsight bias at an aggregate level; respondents recalled having assigned significantly lower probabilities to each event than they actually had (Recall 2001 < Prediction 2001), in line with seeing a safer world but not realizing it (see Table 2 for statistical tests). That occurred despite respondents having reviewed material from the previous year, which could have reinstated cues to their previous perceptions.

As seen in Table 3, respondents' personal experiences varied in ways that were systematically related to their risk judgments. We used these reports to assess hindsight bias at the individual level assuming that personal experience with an event captures part of how respondents view the circumstances determining its probability. We also used Update 2001 judgments as a more encompassing measure of subsequent beliefs.

Personal experience as a measure of hindsight. In 26.0% of all cases (CI: 24.6%, 27.4%), Recall 2001 and Prediction 2001 judgments were identical. Respondents may have either recalled their original judgment or managed to reconstruct the perspective that produced it. In the former case, there is no chance for bias; in the latter case, there is no bias. In 16.2% of all cases (CI: 15.1%, 17.4%), the two judgments differed and respondents reported the experience. In these cases, hindsight bias appeared in respondents recalling a higher probability than the one originally given. It was seen in 40.9% of these comparisons (CI: 37.1%, 44.8%). In 57.8% of all cases (CI: 56.2%, 59.3%), the two judgments differed and respondents did not report the experience. In these cases, hindsight bias appeared in respondents recalling a lower probability than the one originally given. It was seen in 65.1% of these comparisons (CI: 63.1%, 67.0%). Because respondents so seldom reported experiencing these events, the overall pattern is hindsight bias, seen in 61.2% of all possible cases (CI: 59.7%, 62.7%).

Table 4 shows the size of these effects, with Table 5 providing statistical tests. In Table 4, the left-hand side compares original and recalled judgments in cases where respondents reported experiencing each event. For the four events reported by more than 1% of respondents (travel, sleep, mail, flu), Recall 2001 means are significantly lower than Prediction 2001 means, the reverse of hindsight bias (overall $d = -0.28$). The right-hand side has cases where respondents reported not experiencing each event. It shows hindsight bias (Recall 2001 $<$ Prediction 2001) for each event (overall $d = 0.57$). Because the right-hand side involved so many more observations than the left, pooling all responses yields a mean hindsight bias of +9.95% ($d = 0.44$). Thus, the overall pattern was hindsight bias with a magnitude similar to that found elsewhere. Nonetheless, a reverse bias was seen in cases involving events that happened, which have typically produced greater hindsight bias than nonevents (Christensen-Szalanski & Wilham, 1991; Fischhoff & Beyth, 1975). We consider possible explanations below.

Update probabilities as the measure of hindsight. The Update 2001 task sought to capture all that respondents had learned during the year, hence the full set of current beliefs that must be undone in order to avoid hindsight bias. In 87.5% of cases (CI: 86.4%, 88.5%), the previous year was seen as having been safer than it originally had seemed (Update 2001 $<$ Prediction 2001). Hindsight bias was seen in 91.8% of those cases (CI: 90.7%, 92.6%); respondents recalled having given lower probabilities than they actually had (Recall 2001 $<$ Prediction 2001). In the relatively few cases where respondents saw the past year as having been riskier than they originally had (Update 2001 $>$ Prediction 2001), hindsight bias emerged 75.4% of the time (CI: 70.3%, 80.4%); respondents recalled having given a higher probability than they actually had (Recall 2001 $>$ Prediction 2001). Overall, hindsight bias was found in 92.6% of cases (CI: 91.3%, 92.3%) where beliefs had changed (Update 2001 $\neq$ Prediction 2001).

Summary. Respondents showed hindsight bias in the majority of comparisons and of the overall magnitude seen elsewhere (Christensen-Szalanski & Wilham, 1991; Guilbault et al., 2004). The bias was more consistent when assessed with Update 2001 judgments, a new measure intended to capture all current beliefs about the period. By either measure (personal experience, postdiction), there was more bias when it meant recalling lower probabilities than had actually been given. Most cases involved such comparisons because most risks seemed smaller than they had a year earlier, producing hindsight bias overall. Nonetheless, an anomalous reversed bias emerged with the (relatively few) events that respondents reported experiencing personally. One possible explanation is that judgments of specific events were affected by general perceptions of terror risk. If so, then seeing a safer world in 2002 tended to reduce recalled probabilities for all events. Such a global hindsight bias would increase the specific hindsight bias for individual events that had not occurred and would reduce (or even reverse) it for those that had.

Emotion

After examining the impacts of the 2002 emotion manipulations on emotion self-reports, we evaluate their predicted effects on risk judgments and hindsight bias.

Emotion self-reports. Respondents' ratings of their emotions on the five fear and five anger scales revealed several patterns (see Table 6):

Table 4
Personal Experience With Events and Hindsight Bias

	Reported experiencing event[a]							Reported not experiencing event[b]						
		Prediction 2001			Recall 2001				Prediction 2001			Recall 2001		
Event	% of respondents	*M*	*Mdn*	*SD*	*M*	*Mdn*	*SD*	% of respondents	*M*	*Mdn*	*SD*	*M*	*Mdn*	*SD*
Being hurt by terror	0.3	58.5	58.5		65.5	65.5		99.7	20.2	10.0	22.4	16.9	5.0	23.1
Traveling less	30.1	57.5	60.0	37.7	49.3	50.0	35.2	69.9	24.0	5.0	32.1	9.0	0.0	18.1
Trouble sleeping	21.8	42.9	30.0	35.3	29.8	20.0	27.4	78.2	18.7	5.0	27.1	6.0	0.0	13.2
Screening mail	33.5	77.4	90.0	32.0	44.5	50.0	36.7	66.5	45.0	50.0	38.4	7.5	0.0	15.3
Anthrax antibiotics	0.7	86.7	80.0	11.6	70.0	70.0		99.3	21.8	5.0	30.1	10.3	1.0	19.6
Getting flu	27.8	67.4	77.3	32.4	55.1	50.0	31.3	72.2	40.3	50.0	28.2	23.0	15.0	22.5
Other violent crime	0.9	32.5	32.5		39.0	60.0	28.8	99.1	22.8	10.0	23.8	17.8	10.0	21.9

Note. Standard deviations are not reported where there were not enough respondents in the sample (i.e., $n = 2$).
[a] Hindsight bias means Recall $>$ Prediction 2001; respondents who reported experiencing the event recall having assigned higher probabilities than was actually the case. [b] Hindsight bias means Recall $<$ Prediction 2001; respondents who reported not experiencing the event recall having assigned lower probabilities than was actually the case.

Table 5
Comparison of Prediction 2001 With Recall 2001 Judgments for Risk Events Based on Personal Experience With Events

	Reported experiencing the event			Reported not experiencing the event		
Event	*df*	*Z*	*d*	*df*	*Z*	*d*
Terror risks						
Being hurt in a terror attack	1	−0.45		561	−4.16**	0.15
Traveling less	162	−2.64*	.22	398	−8.50**	0.58
Trouble sleeping	121	−4.12**	.42	442	−10.19**	0.60
Screening mail	182	−8.07**	.96	376	−14.50**	1.28
Taking antibiotics against anthrax	3	−1.23		559	−8.82**	0.45
Routine risks						
Getting the flu	154	−3.74**	.39	409	−9.80**	0.68
Being a victim of violent crime (other than terror)	4	−1.04		559	−4.37**	0.22

Note. Asterisks compare Prediction 2001 with Recall 2001 (using Wilcoxon matched, pairs signed-rank tests, given the skewed distributions). Cohen's *d* is not reported when one of the comparison groups is missing a standard deviation that is used to calculate Cohen's *d*.
* $p < .05$. ** $p < .001$.

1. The 2001 emotion manipulations had no carryover effect to 2002. Neutral respondents who had been in the fear and the anger conditions in 2001 reported similar fear, $t(187) = -.22$, *ns*, $d = 0.03$, and similar anger, $t(188) = -.24$, *ns*, $d = 0.03$, in 2002 (when they received no emotion induction). In 2002, they also reported less of the primed emotion than they had in 2001, $t_{fear}(187) = 12.41$, $p < .001$, $d = 0.70$; $t_{anger}(187) = 15.40$, $p < .001$, $d = 1.27$. Thus, any effects of emotion on judgments reflect the inductions received in the 2002 relive condition. These results also show the (desired) failure of the neutral task to evoke these emotions.

2. The 2002 emotion manipulations had the intended effects. Respondents receiving the fear manipulation reported significantly more fear than did those receiving the anger manipulation, $t(198) = -2.57$, $p < .05$, $d = 0.37$, or those in the neutral group, who received no manipulation at all, $t(192) = 4.82$, $p < .001$, $d = 0.64$. Respondents receiving the anger manipulation reported more anger than those receiving the fear manipulation, $t(198) = 4.61$, $p < .001$, $d = 0.65$, or those in the neutral group, $t(194) = 7.25$, $p < .001$, $d = 0.98$.

3. Although the emotion manipulation effectively primed the focal emotion at both times, relive respondents reported less of their primed emotion in 2002 than they had in 2001, $t_{fear}(273) = 7.52$, $p < .001$, $d = 0.55$; $t_{anger}(273) = 5.36$, $p < .001$, $d = 0.29$. Thus, greater distance from September 11th may have reduced respondents' overall emotional responsiveness.

Emotion effects on risk judgments. Figure 1 shows the effects of the emotion manipulations on judgments of terror risks (A) and routine risks (B). Prediction 2002 judgments replicated the pattern of Prediction 2001 judgments (Lerner et al., 2003). Relive respondents receiving the anger manipulation in 2002 saw the five terror risks as significantly less likely than did those in the fear condition, $F(1, 194) = 4.15$, $p < .05$, $MSE = 227.09$, Cohen's $f = .14$. The two groups judged the three routine risks similarly, $F(1, 194) = 0.60$, *ns*, $MSE = 351.39$, Cohen's $f = .04$. Neutral respondents judged the terror risks, $F(1, 177) = 0.20$, *ns*, $MSE = 440.09$, Cohen's $f = .06$, and the routine risks similarly, $F(1, 177) = 0.86$, *ns*, $MSE = 471.83$, Cohen's $f = .03$, regardless of their emotion manipulation in 2001 (as would be expected, given their instruction in 2002 to be unemotional and given the lack of an emotion carryover effect from 2001). Thus, the emotion manipulations, the impacts of which were demonstrated in the self-reports, had their predicted effects on these risk judgments.

There was no Relive × Neutral main effect, terror risks, $F(1, 373) = 0.12$, *ns*, $MSE = 329.71$, Cohen's $f = .04$; routine risks, $F(1, 373) = 1.15$, *ns*, $MSE = 408.31$, Cohen's $f = .07$. Thus, the anger induction reduced risk judgments as much as the fear induc-

Table 6
Emotion Manipulation Checks for Relive and Neutral Respondents in 2001 and 2002 (Mean Ratings)

	Relive				Neutral[a]			
	Fear condition		Anger condition		Fear condition		Anger condition	
Time of report	Fear	Anger	Fear	Anger	Fear	Anger	Fear	Anger
2001	4.57	3.74	2.93	5.25	4.15	3.46	3.06	5.12
2002	3.40	2.99	2.49	4.62	1.93	1.50	1.86	1.57

Note. Respondents in both groups (relive, neutral) were randomly assigned to one of two conditions (fear, anger) and asked to report on the degree of each emotion (fear, anger) that they experienced during the emotion manipulation.
[a] Neutral respondents received the primed emotion in 2001 but not in 2002.

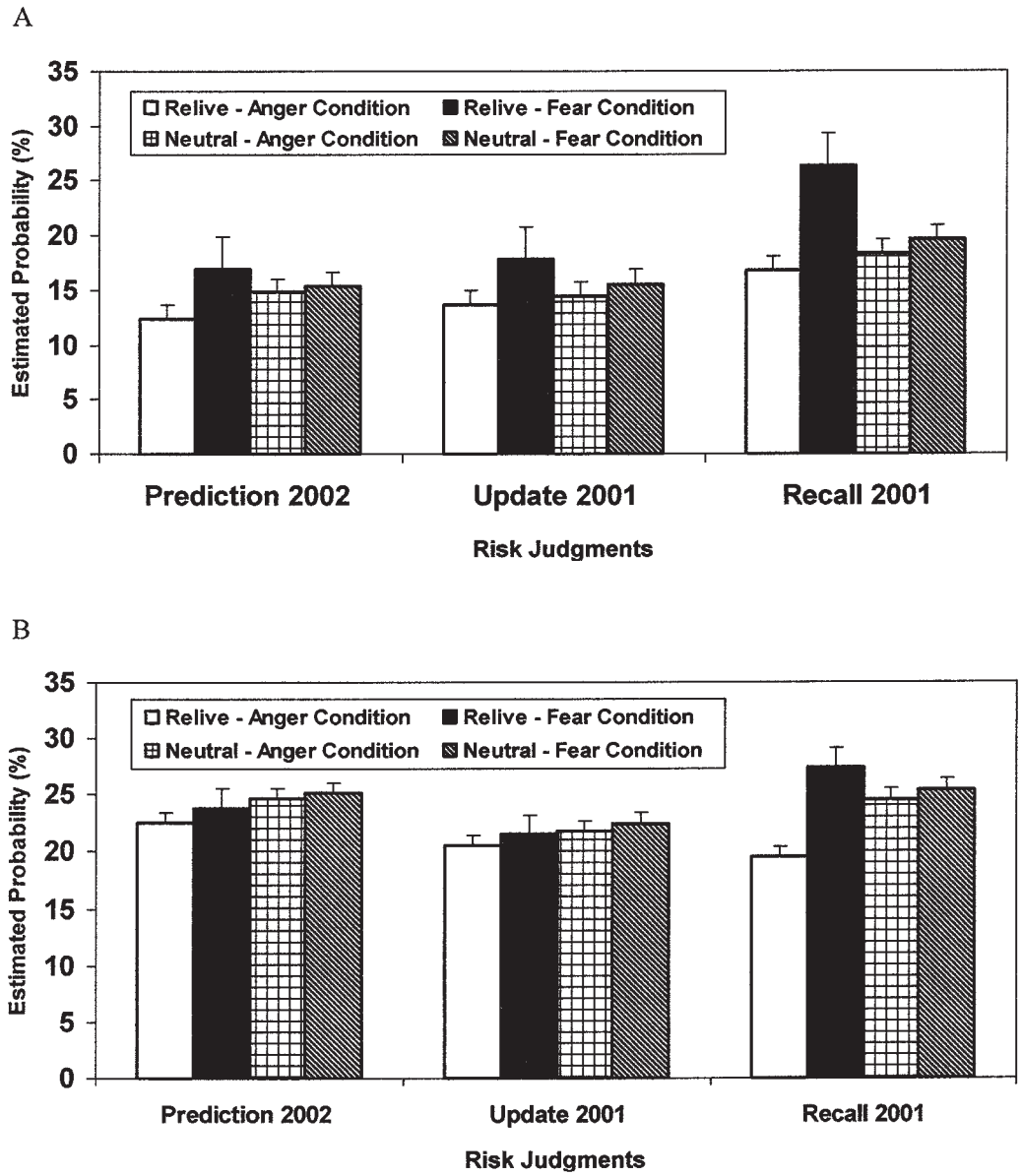

Figure 1. A: Prediction 2002, Update 2002, and Recall 2001 (+ *SE*) for the terror risks judgments for relive and neutral respondents. B: Prediction 2002, Update 2001, and Recall 2001 (+ *SE*) for the routine risk judgments for relive and neutral respondents. For relive respondents, the emotion condition refers to the manipulation that they received in 2001 and 2002. For neutral respondents, the emotion condition refers to the manipulation that they received in 2001.

tion increased them—relative to the baseline provided by the neutral condition.

Update 2001 judgments showed similar emotion induction effects. Respondents receiving the anger manipulation saw the terror risks as having been significantly lower than did those in the fear condition, $F(1, 194) = 4.33$, $p < .05$, $MSE = 277.69$, Cohen's $f = .14$. The two groups judged the routine risks similarly, $F(1, 194) = 2.89$, *ns*, $MSE = 317.54$, Cohen's $f = .09$. Thus, the emotion inductions affected judgments of the past, even for people who had lived through the period—the first evidence of such effects. Neutral respondents saw the previous year similarly, whichever manipulation they had received a year earlier, $F_{terror}(1, 177) = .16$, *ns*, $MSE = 317.13$, Cohen's $f = .06$; $F_{routine}(1, 177) = .55$, *ns*, $MSE = 456.87$, Cohen's $f = .05$. As with Prediction 2002, there was no Relive × Neutral main effect, terror risks, $F(1, 373) = 0.06$, *ns*, $MSE = 297.41$, Cohen's $f = .05$; routine risks, $F(1, 373) = 1.26$, *ns*, $MSE = 385.05$, Cohen's $f = .05$. Thus, the fear and anger inductions had equal and opposite effects on judgments.

The emotion inductions' effects on judgments of the past were also seen in the Recall 2001 responses. Respondents receiving the anger induction recalled having seen significantly smaller terror risks than did those receiving the fear induction, $F(1, 194) =$

11.88, $p < .001$, $MSE = 314.58$, Cohen's $f = .26$. They also recalled having seen significantly lower routine risks, $F(1, 194) = 9.29$, $p < .001$, $MSE = 324.49$, Cohen's $f = .21$. Neutral respondents recalled having given similar probabilities, whichever manipulation they had received a year earlier, $F_{terror}(1, 177) = .07$, *ns*, $MSE = 404.53$, Cohen's $f = .07$; $F_{routine}(1, 177) = .79$, *ns*, $MSE = 368.58$, Cohen's $f = .03$. There was no Relive × Neutral main effect, terror risks, $F(1, 373) = 0.41$, *ns*, $MSE = 366.58$, Cohen's $f = .03$; routine risks, $F(1, 373) = 0.81$, *ns*, $MSE = 352.83$, Cohen's $f = .03$. Thus, the anger induction decreased recalled predictions as much as the fear induction increased them.

Summary. As in Lerner et al. (2003), the emotion inductions affected reported emotion experiences and, with them, judgments of future risks—although the effects were smaller than in 2001 ($d = 0.14$ vs. 0.26). As predicted by appraisal-tendency theory, the emotion inductions also affected how respondents reevaluated and recalled the past ($d = 0.14$, 0.26, for Update 2001 and Recall 2001, respectively). These novel effects were observed despite respondents having concrete cues about that past (of a type found to suppress incidental emotion effects; Forgas, 1995). Compared with a neutral baseline, the anger manipulation reduced perceptions of risk as much as the fear manipulation increased them, consistent with the equal effects seen in the emotion self-reports. These effects were found despite the inductions raising the target emotions to lower levels than they had in 2001.

Emotion and Hindsight Bias

The emotion manipulations sought to reinstate relive respondents' experience of a year earlier—beyond the memories evoked by neutral respondents' review of the previous year's materials. The somatic markers associated with these recreated emotions might provide cues to prior beliefs, in addition to those found in the content of the inductions (written passages, news pictures, and audio clips). As a result, we expected the appraisal tendencies associated with fear and anger to shape memories and judgments in opposite ways—just as they had a year earlier.

Debiasing failed. Recall 2001 judgments were equally far from the Prediction 2001 judgments for relive and neutral respondents for all eight risks, $t(381) = -1.44$, *ns*, $d = 0.04$, for the five terror risks, $t(381) = 1.78$, *ns*, $d = 0.05$, and for the three routine risks, $t(198) = 0.98$, *ns*, $d = 0.04$. Debiasing failed even though the emotion inductions had other effects. Possibly, debiasing would have been observed with a stronger manipulation, more fully recreating the emotional content and cues associated with the respondents' writing and reflections or with a neutral condition that evoked none of the previous year's perspective.

Discussion

Theoretical Contributions

We have used the natural experiment of September 11th to test several behavioral theories with a study having several features of ecological validity. It focused on salient events with common tasks (prediction, recall) administered in respondents' homes. It used a random sample of Americans, a population seldom available for psychological research. We found that several theories established in laboratory settings also operate in this context, indirectly supporting the ecological validity of other experimental studies of judgment, memory, and affect. The relative verisimilitude of the tasks afforded effect-size estimates in contexts where one can begin thinking about their applied importance—which will, of course, depend on the sensitivity of specific decisions to the values of specific judgments. The study also asked new questions about these basic behavioral processes and the interactions between them, revealing several new patterns. After reviewing the main results, we consider possible limitations and applied implications.

Availability-based judgment. Comparing predictions from 2001 and 2002 revealed a decrease in the perceived threat of terror consistent with respondents' few reported experiences with most of these events. These reports were correlated with respondents' risk judgments in ways that suggest orderly inferential processes; in 2001, respondents made predictions that anticipated these experiences, showing insight into their relative risk levels (possibly inflated by those already experiencing or intending an act). In 2002, respondents' experiences were correlated with three sets of judgments: (a) memories of their previous judgments (Recall 2001), (b) postdicted assessments of risk (Update 2001), and (c) predictions for the coming year (Prediction 2002). For each judgment, correlations were stronger for more direct experiences. Indeed, only respondents' personal experience had an independent contribution to predicting their probability judgments—as though they inferred little from others' experience, even for rare events, where an individual's personal experience says little.

These results are consistent with many memory- and experience-related processes. However, a parsimonious account is that people rely on the availability heuristic, operationalized as direct experience, when making probability judgments. The correlations with Prediction 2002 judgments extend Lichtenstein et al.'s (1978) between-respondent result to a within-respondent design. The other relationships are new results; postdictions have not previously been studied. Alternative measures of availability have not been related to how people recall past predictions, how they revise those predictions in hindsight, or how accurately they make them.

Hindsight. The difference between risk judgments in 2001 and 2002 created the possibility of hindsight bias. The correlations between these judgments and reported experiences suggest that respondents may have reworked their beliefs in ways that would be difficult to undo, increasing the risk of hindsight bias. That bias was, indeed, found; when asked to recall their 2001 predictions, respondents exaggerated how much those beliefs had resembled their present ones. The overall pattern (see Table 1) was to recall having seen the world as safer than it had, in fact, seemed—looking backward from a safer-seeming world. The bias occurred despite respondents having reviewed materials from the previous year.

Item-level analyses used two measures of belief change. The first was respondents' direct experience, the report most strongly associated with their risk judgments. As mentioned, respondents typically reported not having experienced these events. In those cases, they typically recalled having given a lower probability than they actually had, representing hindsight bias. However, they also underestimated the probabilities that they had assigned to the (relatively few) events that they reported having experienced, reversing the bias. Nonevents were six times more common than events. As a result, there were many more cases of bias than of reverse bias so that hindsight bias predominated. The mean overall difference of 10% between actual and recalled predictions, in the direction of hindsight bias, resembles that found in Christensen-

Szalanski and Wilham's (1991) review. The overall effect size ($d = 0.31$) is consistent with that found in Guilbault et al.'s (2004) review.

The second item-level analysis used a new measure of belief change: postdiction, a task that had respondents reevaluate the preceding year's risks, considering all that they now knew, including their personal experiences. This broader measure revealed more widespread hindsight bias, both for events seen as having been safer than they originally had and for (the relatively few) events seen as having been riskier. The bias was more common in the former case, similar to its greater strength with nonevents. Consistency checks provided some evidence that respondents understood the task (i.e., they did not see what happened as having been inevitable).

The differential effect with events and nonevents reverses the usual pattern of greater hindsight bias with events (Christensen-Szalanski & Wilham, 1991; Fischhoff & Beyth, 1975). The usual finding has been attributed to events evoking deeper processing than nonevents, making their effects harder to undo. However, in the year following September 11th and the anthrax crisis, the big news might have been how few additional terror-related events happened, as seen in respondents' overestimation of most events' likelihood. If so, then nonevents could have had greater cognitive impact, thereby producing greater hindsight bias.

However, the bias was not just weaker when assessed in terms of events that respondents experienced, but reversed. A speculative account is that specific risks are judged in the context of general perceptions of risk. Seeing a generally safer world would, then, reduce all risk judgments. Not realizing that beliefs had changed in this general way would reduce recalled probabilities for all events. Such a global hindsight bias would enhance the specific bias for nonevents, but diminish it for events, perhaps even reversing it. From this perspective, beliefs about individual risks must be considered in the context of beliefs about the general situation (Fischhoff, 2002; Fischhoff, Gonzalez, Small, & Lerner, 2003a).

Emotion. These results extend understanding of the interplay of emotion and judgment in three ways predicted by appraisal-tendency theory, which posits broad, emotion-specific effects on judgment (Lerner & Keltner, 2000, 2001):

1. Incidental emotions can color not only judgments of an uncertain future (Prediction 2001 and 2002) but also judgments of a concrete past. Specifically, the emotion manipulations affected both memories of risk judgments (Recall 2001) and judgments of what risks had been (Update 2001), even though respondents presumably had had extensive exposure to real-world attention to these events. Thus, even incidentally primed emotions can serve as perceptual lenses affecting perceptions of the future, present, and even one's own past.

2. Although self-reported emotional levels had declined over the year, the manipulations still affected the focal emotions and, through them, risk judgments. This study was the first to track the impact of emotion primes over a long time period (one year) and as function of distance from a highly salient emotional target event (September 11th).

3. Having a neutral condition in 2002 allowed assessing the absolute (and not just relative) impact of the emotion inductions. We found no main effect on risk perceptions, meaning that the inductions did not generally increase or decrease risk judgments. Rather, the fear manipulation increased risk judgments as much as the anger manipulation reduced them.

Emotion and hindsight. Reliving the 2001 emotion manipulation was intended to help reinstate respondents' earlier feelings. If it also evoked previous memories and appraisals, then it could reduce hindsight bias. However, despite affecting reported emotions and risk judgments, this way of revisiting the past did not reduce hindsight bias. These results suggest a core of risk beliefs that were not affected by the manipulation along with a more malleable periphery of beliefs that were. This manipulation did not bring emotions back to their original intensity. Possibly, a more powerful emotion manipulation would shrink the core and reduce the extent of hindsight bias. A more "neutral" neutral condition might show a larger comparative difference.

Methodological considerations. Probability judgments have often been found to have good reliability and some forms of validity (e.g., Wallsten & Budescu, 1983; Woloshin, Schwartz, Byram, Fischhoff, & Welch, 2000; Yates, 1990), at least for the ordinal scale uses that predominated here. We sought to use relatively well-defined events whose occurrence could be assessed with respondents' own reports (Fischhoff, 1994). Nonetheless, some caution is warranted in taking probability judgments as expressing absolute expectations. For example, the disparity between means and medians reflected skewed distributions. High values might have come from both respondents who saw large risks and from ones who used *50* as a form of "do not know" (in the sense of *50/50*; e.g., Fischhoff & Bruine de Bruin, 1999; Fischhoff et al., 2000). The predictive validity of the Prediction 2001 judgments might have been reduced by such anomalous (50/50) responses. It might have been inflated by respondents who were already experiencing an event or intending to do it (e.g., reducing travel, losing sleep), recognizing that intentions are imperfect predictors of behavior (Morwitz, 2001).

Although prediction and recall are familiar tasks, postdiction might not be. The Update 2001 responses showed some construct validity with respondents updating their judgments in keeping with their experiences and rejecting naive determinism (i.e., not assigning 100% to events that had happened and 0% to those that had not). Our conclusions stand even if the postdictions are ignored. Nonetheless, they are, in principle, a more appropriate measure of belief change by virtue of encompassing more that just direct experience.

With any repeated measures design, earlier tasks might influence later ones. A year separated our two data collections, which should minimize any direct influence as seen in the absence of emotion carry-over effects. Recall 2001 responses equaled their Prediction 2001 target about 25% of the time. There is no way to know which of these matches reflect actual memory, coincidental reuse of a response, or success at reconstructing past perspectives.

In 2002, Recall 2001 was the first task, in order to provide the clearest test of hindsight bias. The lower recalled probabilities presumably reflected the perception of a safer world as seen in the Update 2001 and Prediction 2002 judgments. Once provided, Recall 2001 responses may have artifactually anchored those that followed. However, that should not have affected our predominantly correlational analyses of those judgments.

As mentioned, a general tendency to give lower responses could have contributed to the differential hindsight bias result with events and nonevents. Lower responses reduce the bias for events

and increase it for nonevents. Also as mentioned, a plausible source of generally lower responses is seeing the world as generally safer. In that context, the few events that occurred are noteworthy in a year that produced much less terror than respondents had feared. Some of those events might even have been seen as surprises, which could have led to overcorrection (e.g., "I cannot imagine myself having predicted screening my mail for anthrax.")

The study's ecological validity inevitably entailed some loss of experimental control. Respondents completed the tasks in their own homes and on their own TVs, without research staff monitoring them. Requiring them to write about their emotions in 2001 was a stronger way of ensuring involvement than asking them to review that text in 2002. Thus, although the 2002 manipulation was strong enough to induce emotion effects, those were perhaps weaker than they might have been with direct experimenter control.

Applied Implications

As mentioned, decisions vary in their sensitivity to changes in inputs (like probability judgments). A 10% difference in risk judgments, due to fear and anger manipulations, might tip some decisions while having little effect on others. Similarly, a 10% bias in recalled probabilities might shift some past decisions from seeming sensible to seeming imprudent while not affecting the evaluations of others. The effect sizes observed here allow conducting sensitivity analyses of specific choices. Sometimes, the same shift may affect several inputs, as when a policy depends on multiple risks, each of which may seem greater or smaller in one perspective than another. For example, a fearful citizen may look everywhere and see somewhat greater risks with a cumulative impact that justifies more aggressive antiterror policies.

In addition to its implications for specific decisions, each study of judgment contributes to our overall understanding of the citizens' role in (public and private) decision making:

1. Citizens' risk judgments respond to their observations in orderly ways. A nationally representative sample made generally defensible inferences when updating their beliefs. They relied most heavily on the most immediate experiences. They recognized the unexpected safety of the year between experiments, while realizing that it had not been guaranteed (Update 2001) and might not be sustained (Prediction 2002). They had some insight into their relative degree of risk (Prediction 2001) as measured by their subsequent experiences. Given citizens' capacity to learn, it is important to provide them with ready access to accurate, relevant information, especially when natural sampling creates an incomplete or distorted picture (Fischhoff, 1999; National Research Council, 1989; U.S. General Accounting Office, 2004). Citizens say that they want to know about terror risks even if the information worries them (Fischhoff et al., 2003a).

2. Expect hindsight bias, even with salient events; do not rely on event-aided memory to overcome it. The bias was observed here, despite involving well-publicized, personally relevant events. It was not reduced by an intervention designed to reinstate prior cognitive and emotional perspectives. As a result, ensuring fair judgment of past actions requires special efforts beyond asking people to focus on how things once looked and felt. Rather, people need explicit retrieved evidence of the sort produced by investigative committees or retrieved from diaries and archives. Where there are no relevant records, it might help to create alternative scenarios for how things might have been different (Fischhoff, 1982). The reduced, and even reversed, bias with reported events suggests that help is especially needed when a specific event belongs to a general class, such that general perceptions affect judgments of the specific event. The magnitude of the bias (about 10% on average) provides a perspective for evaluating claims (e.g., by political leaders) of having been judged unfairly in hindsight.

3. People need protection from emotional manipulation, even (or especially) with fateful topics. Stimuli drawn from the national news media, supported by personal reflection, affected respondents' emotional states enough to influence their risk judgments. That happened even a year after the events of September 11th, during which many related events had doubtless been seen and experienced. Emotions affected judgments not only when respondents judged their future but also when they judged their past. These emotion effects could be large enough to tip the scales in close decisions. Thus, citizens are at risk of manipulation—being made angry enough to accept aggressive policies (which seem more likely to succeed) or being made anxious enough to accept precautionary ones.

References

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* New York: Cambridge University Press.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997, February 28). Deciding advantageously before knowing the advantageous situation. *Science, 275,* 1293–1294.

Beyth-Marom, R., & Fischhoff, B. (1977). Direct measures of availability and judgments of category frequency. *Bulletin of the Psychonomic Society, 9,* 236–238.

Bower, G. H. (1982). Mood and memory. *American Psychologist, 36,* 129–148.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty, 19,* 7–48.

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

Combs, B., & Slovic, P. (1979). Newspaper coverage of causes of death. *Journalism Quarterly,* 56, 837–843, 849.

Damasio, A. R. (1995). *Descartes' error: Emotion, reason, and the human brain.* New York: Avon Books.

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

Denis, J. M. (2001). Are internet panels creating professional respondents? *Marketing Research,* Summer, 34–38.

Fischhoff, B. (1975). Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 104,* 288–299.

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

Fischhoff, B. (1994). What forecasts (seem to) mean. *International Journal of Forecasting, 10,* 387–403.

Fischhoff, B. (1995). Risk perception and communication unplugged: Twenty years of process. *Risk Analysis, 15,* 137–145.

Fischhoff, B. (1999). Why (cancer) risk communication can be hard. *Journal of the National Cancer Institute Monographs, 25,* 7–13.

Fischhoff, B. (2002). Assessing and communicating the risks of terrorism. In A. H. Teich, S. D. Nelson, & S. J. Lita (Eds.), *Science and technology*

in a vulnerable world (pp. 51–64). Washington, DC: American Association for the Advancement of Science.

Fischhoff, B. (in press). Risk perception and communication. In D. Kamien (Ed.), *McGraw-Hill handbook of terrorism and counter-terrorism.* New York: McGraw-Hill.

Fischhoff, B., & Beyth, R. (1975). "I knew it would happen"– Remembered probabilities of once-future things. *Organizational Behavior and Human Performance, 13,* 1–16.

Fischhoff, B., Bostrom, A., & Quadrel, M. J. (2002). Risk perception and communication. In R. Detels, J. McEwen, R. Beaglehole, & H. Tanaka (Eds.), *Oxford Textbook of Public Health* (pp. 1105–1123). London: Oxford University Press.

Fischhoff, B., & Bruine de Bruin, W. (1999). Fifty/fifty = 50? *Journal of Behavioral Decision Making, 12,* 149–163.

Fischhoff, B., Bruine de Bruin, W., Perrin, W., & Downs, J. S. (2004). Travel risks in a time of terror. *Risk Analysis, 24,* 1299–1307.

Fischhoff, B., Gonzalez, R. M., Small, D. A., & Lerner, J. S. (2003a). Evaluating the success of terror risk communication. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science, 1,* 255–258.

Fischhoff, B., Gonzalez, R. M., Small, D. A., & Lerner, J. S. (2003b). Judged terror and distance from the World Trade Center. *Journal of Risk and Uncertainty, 26,* 137–151.

Fischhoff, B., Parker, A., Bruine de Bruin, W., Downs, J., Palmgren, C., Dawes, R. M., et al. (2000). Teen expectations for significant life events. *Public Opinion Quarterly, 64,* 189–205.

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin, 117,* 39–66.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of judgment.* New York: Cambridge University Press.

Gray, J. R. (2004). Integration of emotion and cognitive control. *Current Directions in Psychological Science, 13,* 46–48.

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

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

Hoffrage, U., Hertwig, R., & Gigerenzer, G. (2000). Hindsight bias: A by-product of knowledge updating? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 566–581.

Krotki, K., & Dennis, J. M. (2001, August). *Probability-based survey research on the Internet.* Paper presented at the 53rd conference of the International Statistical Institute, Seoul, Korea.

Leiss, W., & Chociolko, C. (1997). *Mad cows and mothers milk.* Kingston & Montreal: Queens and McGill University Press.

Lerner, J. S., & Gonzalez, R. M. (2004). *On feeling angry and self assured*. Paper presented at the Self and Identity Preconference at the annual meeting of the Society for Personality and Social Psychology, Austin, TX.

Lerner, J. S., & Gonzalez, R. M. (2005). Forecasting one's future based on fleeting subjective experiences. *Personality and Social Psychology Bulletin.*

Lerner, J. S., Gonzalez, R. M., Small, D. A., & Fischhoff, B. (2003). Emotion and perceived risks of terrorism: A national field experiment. *Psychological Science, 14,* 144–150.

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgment and choice. *Cognition and Emotion, 14,* 473–493.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology, 81,* 146–159.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin, 125,* 255–275.

Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 551–578.

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes, 65,* 272–292.

Loewenstein, G., & Adler, D. (1995). A bias in the prediction of tastes. *Economic Journal: The Journal of the Royal Economic Society, 105,* 929–937.

Loewenstein, G., & Lerner, J. S. (2003). The role of affect in decision making. In R. Davidson, H. Goldsmith, & K. Scherer (Eds.), *Handbook of affective science* (pp. 619–642). Oxford, United Kingdom: Oxford University Press.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, E. (2001). Risk as feelings. *Psychological Bulletin, 127,* 267–286.

Lupia, A. (1994). Shortcuts versus encyclopedias. *American Political Science Review, 88,* 63–76.

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

Mazursky, D., & Ofir, C. (1990). "I could not 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.

Mellers, B., Schwartz, A., & Ritov, I. (1999). Emotion-based choice. *Journal of Experimental Psychology: General, 128,* 332–345.

Morwitz, V. G. (2001). Methods for forecasting from intentions data. In J. S. Armstrong (Ed.), *Principles of forecasting* (pp. 33–56). Norwell, MA: Kluwer.

National Research Council. (1989). *Improving risk communication.* Washington, DC: National Academy Press.

Reyna, V. F. (2004). How people make decisions that involve risk: A dual-processes approach. *Current Directions in Psychological Science, 13,* 60–66.

Reyna, V. F., & Brainerd, C. J. (1991). Fuzzy-trace theory and framing effects in choice: Gist extraction, truncation, and conversations. *Journal of Behavioral Decision Making, 4,* 249–262.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7,* 1–75.

Rich, F. (2004, June 27). The best Goebbels of all? *New York Times,* Section 2, pp. 1, 12.

Schwarz, N., & Clore, G. (1996). Feelings and phenomenal experiences. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 103–119). New York: The Guilford Press.

Schwarz, N., & Vaughan, L. A. (2002). The availability heuristic revisited: Ease of recall and content of recall as distinct sources of information. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 397–420). Cambridge, England: Cambridge University Press.

Slovic, P. (Ed.). (2001). *Perception of risk.* London: Earthscan.

Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). The affect heuristic. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: Psychology of intuitive judgment* (pp. 397–420). Cambridge, England: Cambridge University Press.

Smith, C. A., & Ellsworth, P. E. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology, 48,* 813–838.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207–232.

U.S. General Accounting Office. (2004). *Risk communication principles may assist in refinement of the Homeland Security Advisory System* (Publication no. GAO-04–538T). Washington, DC: U.S. General Accounting Office.

Van Boven, L., Dunning, D., & Loewenstein, G. (2000). Egocentric

empathy gaps between owners and buyers: Misperceptions of the endowment effect. *Journal of Personality and Social Psychology, 79,* 66–76.

Viscusi, W. K., & O'Connor, C. J. (1984). Adaptive responses to chemical labeling: Are workers Bayesian decision makers? *American Economic Review, 74,* 942–956.

von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research.* New York: Cambridge University Press.

Walker, D. A. (2003). Converting Kendall's Tau for correlational or meta-analytic analyses. *Journal of Modern Applied Statistics, 2,* 525–530.

Wallsten, T. S., & Budescu, D. V. (1983). Encoding subjective probabilities: A psychological evaluation and psychometric review. *Management Science, 29,* 151–173.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction. *Psychological Bulletin, 116,* 117–142.

Wilson, T. D., & Gilbert, D. T. (2003). Affective forecasting. *Advances in Experimental Social Psychology, 35,* 345–411.

Woloshin, S., Schwartz, L. M., Byram, S., Fischhoff, B., & Welch, H. G. (2000). A new scale for assessing perceptions of chance: A validation study. *Medical Decision Making, 20,* 298–307.

Yates, J. F. (1990). *Judgment and decision making.* Englewood Cliffs, NJ: Prentice Hall.

Received October 28, 2004
Revision received March 15, 2005
Accepted March 15, 2005 ■

Correction to Fischhoff et al. (2005)

The following article from the June 2005 issue has been corrected:

Fischhoff, B., Gonzalez, R.M., Lerner, J.S., & Small, D.A. (2005). Evolving judgments of terror risks: Foresight, hindsight, and emotion. *Journal of Experimental Psychology: Applied, 11,* 124-139.

After an internal University investigation and independent review by the federal agencies that funded her research, it was found that the second author on the original paper, Roxana M. Gonzalez, falsified data in connection with certain research projects.

Although Ms. Gonzalez did not admit to any acts of falsification of data in connection with this particular project and paper, the University conducted a comprehensive review of all published papers on which Ms. Gonzalez worked. In this regard, researchers not involved in this original project (Wändi Bruine de Bruin and Lei Lai) were engaged to repeat all analyses reported in the article, starting with the original, untainted data for the study. Based on this reanalysis, a corrected version of the article is now being published. The corrected and republished version is available electronically by going to http://dx.doi.org/10.1037/a0027959

Roxana M. Gonzalez has taken full responsibility for her acts of research misconduct, and has acknowledged that her coauthors were not aware of her actions.

DOI: 10.1037/a0027958