



ACADEMIC
PRESS

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Organizational Behavior and Human Decision Processes 90 (2003) 244–261

ORGANIZATIONAL
BEHAVIOR
AND HUMAN
DECISION PROCESSES

www.elsevier.com/locate/obhdp

Betrayal aversion: When agents of protection become agents of harm

Jonathan J. Koehler^{a,*} and Andrew D. Gershoff^b

^a *McCombs School of Business and School of Law, The University of Texas at Austin, CBA 5.202, Austin, TX 78712-1175, USA*

^b *Department of Marketing, Columbia University Graduate School of Business, 520 Uris Hall, 3022 Broadway, New York, NY 10027-6902, USA*

Abstract

A form of betrayal occurs when agents of protection cause the very harm that they are entrusted to guard against. Examples include the military leader who commits treason and the exploding automobile air bag. We conducted five studies that examined how people respond to criminal betrayals, safety product betrayals, and the risk of future betrayal by safety products. We found that people reacted more strongly (in terms of punishment assigned and negative emotions felt) to acts of betrayal than to identical bad acts that do not violate a duty or promise to protect. We also found that, when faced with a choice among pairs of safety devices (air bags, smoke alarms, and vaccines), most people preferred inferior options (in terms of risk exposure) to options that included a slim (0.01%) risk of betrayal. However, when the betrayal risk was replaced by an equivalent non-betrayal risk, the choice pattern was reversed. Apparently, people are willing to incur greater risks of the very harm they seek protection from to avoid the mere possibility of betrayal.

© 2002 Elsevier Science (USA). All rights reserved.

Keywords: Betrayal; Trust; Promise; Risk; Air bags

1. Trust and betrayal

Trust plays a central role in personal and business relationships (Barber, 1983; Mayer, Davis, & Schoorman, 1995). A recent cross-disciplinary review defined trust as “a psychological state composing the intention to accept vulnerability based on positive expectations of the intentions or behavior of another” (Rousseau, Sitkin, Burt, & Camerer, 1998, p. 395). This definition indicates that there are three components at the centerpiece of a trusting relationship: (a) dependency among parties, (b) vulnerability of at least one party, and (c) positive (i.e., confident) expectations or beliefs by at least one party.

Some trust relationships are asymmetric in the sense that the dependency, vulnerability and confident expectations loom larger for one party than the other. Consider, for example, the trust between those who have a professional duty to protect and those who receive the protection. Military leaders protect our country. Firefighters protect our homes. Apartment security guards

protect our families. Nannies protect our children. Nurses and doctors protect our bodies. Psychiatrists, and counselors protect our minds. In all cases, there is an asymmetric trust flowing from the protectees to the protectors.

This paper examines how people respond to violations and anticipated violations of asymmetric protective trusts. We also investigate specific mechanisms that may be responsible for why people respond to such violations as they do. Our focus throughout is on a subset of asymmetric protective trust violations in which the party that is obligated to protect causes the very harm that he or she was entrusted to guard against. Examples include the military general who passes secrets to the enemy and the security guard who robs the jewelry store where he is employed. We refer to such violations of trust as “betrayals.” We believe that these types of betrayals are among the most harmful and are particularly likely to be resisted and punished.

1.1. What is betrayal?

Betrayal is sometimes defined as a violation of “pivotal” expectations of a trustor (Elangovan & Shapiro, 1998; Morris & Moberg, 1994). Similarly, the dictionary

* Corresponding author. Fax: 1-512-471-0587.

E-mail addresses: koehler@mail.utexas.edu (J.J. Koehler), ag644@columbia.edu (A.D. Gershoff).

defines betrayal as “a violation of trust or confidence, an abandonment of something committed to one’s charge” (Oxford English Dictionary, 1989). An implication of these definitions is that betrayals may result from actions and inactions alike. We accept this point but, at this early stage, we confine ourselves to the type of betrayals in which an agent of protection *actively causes* the very harm he or she was entrusted to prevent. More specifically, we focus on how people *respond* to such betrayals.

1.2. Response to betrayal

Victims of such classic betrayals as marital infidelity and incest respond in various ways (Davis & Petretic-Jackson, 2000). Some try to disentangle themselves from the betrayer (Buunk, 1982), some punish the betrayer, and some do not respond at all. Although there is no uniform behavioral response to classic betrayals, fight or flight behaviors are common.

The emotional experience associated with betrayal typically includes visceral, intense, and protracted negative feelings (Finkelhor & Browne, 1985; Forward & Bruck, 1978; Strauss, 1994). For example, children who are abused or otherwise betrayed by adults commonly experience an “unrelenting anger” and view the world as a disordered place (Larson, 1993).

Though the feelings may be less intense, victims of betrayals in the workplace show a similar pattern. In a series of provocative studies, Robinson and her colleagues found that betrayals of “psychological contracts” in employment settings reduced employees’ satisfaction, trust, and intent to stay with the company (Morrison & Robinson, 1997; Robinson, 1996; Robinson & Rousseau, 1994). A recent longitudinal study of perceived psychological contract breach found that the most intense feelings of violation occurred among employees who believed that their employers purposefully renege on their promises (Robinson & Morrison, 2000). Similarly, Medvec, Valley, and Thaler (1999) observed a “concession aversion” in which mock employees who perceived a breach in the employer–employee relationship were willing to leave the betraying employer even when economic considerations favored staying with the company. In short, there is increasing evidence that betrayals of various types produce strong, negative feelings. Victims frequently respond by punishing or distancing themselves from the responsible agents.

1.3. Can objects betray?

Betrayals are ordinarily accompanied by the perception of intentionality (Akerstrom, 1991). Those who do not intend to betray or who are unaware of their role as trusted partner may not be viewed as betrayers even

when they cause harm (but see Elangovan & Shapiro, 1998; for commentary on “accidental betrayals”). Because inanimate objects are incapable of intentionality or awareness, they cannot *really* betray our trust. Nevertheless, people do form asymmetric trusting relationships with various products and objects. For example, we trust foods, medicines, and safety devices to preserve our health and to protect us from injury or death. As in cases involving trusted human agents, these products *seem* to “betray” us when they cause the very harms they were designed to guard against. Recent examples of safety product betrayals include: an asthma medication that causes airway constriction (Asmus, Barros, Liang, Chesrown, & Hendeles, 2001); a sunscreen that causes skin cancer (Sunscreen takes, 1998); a bear repellent that attracts bears (Matthews, 1999) and a water filtration system that contaminates water (Barnum, 1998). How do people respond to such apparent object betrayals?

The workplace studies identified above notwithstanding, research on response to human betrayal, object betrayal, and the mere *possibility* of betrayal is in its infancy. To date, no studies have systematically examined: (a) the desire to punish betrayals after they occur, (b) the psychological mechanism(s) that produce the betrayal response, (c) whether people respond to object betrayals as they do to more conventional betrayals, (d) variables that may mediate the negative feelings that betrayals evoke, or (e) whether people are betrayal averse when choosing among options. We present five studies that address these issues and identify a pattern of responses that we call betrayal aversion. We find theoretical support for betrayal aversion in the trust literature and in the literature on the impact of affect and other non-rationalist influences on risky decision-making. We conclude that the desire to avoid betrayal risks and to punish betrayals when they occur reflects the importance of maintaining trust and order in the social world.

2. Study 1: Betrayal punishments

Consider once again the professional protectors mentioned above. How might we react if they not only failed to protect us, but betrayed us by causing the very harm they were supposed to guard against? Imagine a firefighter who commits arson, a nanny who shakes a baby to death, or a nurse who injects lethal substances into healthy patients. Such criminal actions warrant punishment regardless of who committed them. But the element of betrayal associated with each would seem to make them even more worthy of punishment.

Some anecdotal evidence suggests that courts are more likely to impose harsh punishments “when they sense a promise betrayed” (Scallen, 1993, p. 904; see also Giles, 1995; Rose, 1995). Examples include attorneys

who take advantage of vulnerable clients (People v. Lowery, 1995), insurance companies that deny claims in bad faith (Baker, 1994, p. 1399), and violations of trust that occur even when there is no actual harm (Snepp v. US, 1980).¹ Still, the claim that there is a relationship between betrayal and punishment is largely untested. Study 1 addressed this concern. We hypothesized that people would impose harsher punishments on people who committed crimes that involve betrayals than on those who committed identical crimes that do not involve betrayals.

2.1. Method

2.1.1. Participants

Fifty-six undergraduates at the University of Texas at Austin and Columbia University participated in Study 1 in exchange for course credit or a payment of \$3.00.

2.1.2. Design and procedure

The experiment was conducted in a classroom setting where participants were randomly assigned to a betrayal or non-betrayal condition and provided with a two-page stimulus. On the first page, participants read brief descriptions of five different crimes and were asked to recommend jail-time punishments (in months or years) for each of the perpetrators. The five offenses (which appeared in random order) were: credit card fraud, bank robbery, treason, child molestation, and rape. We operationalized the between-subjects betrayal variable by varying the profession of the perpetrators. All five perpetrators in the betrayal condition caused harms that they would be expected to guard against by virtue of their profession. No such professional duty to protect was present in the non-betrayal condition. The respective professions in the betrayal and non-betrayal conditions for each of the five crimes were: (a) telephone salesperson/administrative assistant (credit card fraud), (b) security guard/janitor (bank robbery), (c) military leader/orchestra conductor (treason), (d) day-care worker/grocery clerk (child molestation), and (e) campus police officer/construction worker (rape).

Participants also rated the prestige of five professions on a scale from 1 (not at all prestigious) to 7 (highly prestigious). The rated professions were those of the perpetrators from the condition that participants were *not* in. This was done to ensure that the prestige ratings were not associated with the punishment recommenda-

¹ In Snepp v. US (1980), the US Supreme Court found that a former CIA agent “irreparably harmed the United States government” (p. 513) when he published a book without clearing its contents with the CIA, as required by the terms of his employment. Although the book did not contain classified or otherwise harmful information, the former agent was forced to pay the government \$144,000 in damages for this violation of trust (Allen, 1999).

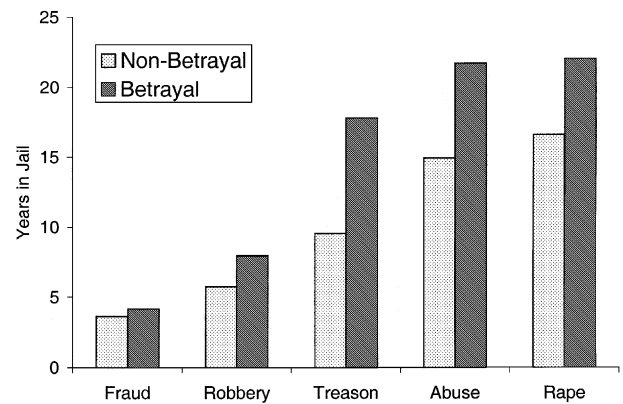


Fig. 1. Study 1: Mean punishment (years in jail) by type of crime.

tions. These ratings served as a check on our assumption that the professions in the betrayal and non-betrayal conditions did not differ in prestige.

A 5×2 mixed design was used in which crime was a five-level within-subjects variable and betrayal was a two-level between-subjects variable. We predicted a main effect for betrayal on amount of perceived betrayal and recommended jail times.

Before running this experiment, we pre-tested the materials on a separate group to ensure successful manipulation of the betrayal variable. Pre-test participants rated the degree of betrayal in either the set of 5 betrayal ($n = 20$) or non-betrayal ($n = 21$) offenses using 7-point Likert-type scales.

2.2. Results and discussion

Our pre-test confirmed that there was more perceived betrayal for those professions that implied a specific duty to protect ($M = 6.3$) than those that did not ($M = 4.5$), $F(1, 39) = 32.44$, $p < .001$. This suggests that perceived betrayal was manipulated successfully across conditions. We also confirmed that perceived job prestige did not differ significantly between the betrayal and non-betrayal conditions, $F(1, 53) = 2.42$, $p > .05$. This rules out the possibility that any observed difference in amount of punishment between conditions was caused by differences in the prestige of the perpetrators.

To examine our central hypothesis, we performed a 5 (crime: credit card fraud, bank robbery, treason, child molestation, rape) $\times 2$ (betrayal: yes, no) repeated measures ANOVA with recommended jail time as the dependent variable.² We observed significant main effects for crime, $F(4, 216) = 31.00$, $p < .001$, and betrayal, $F(1, 54) = 4.24$, $p < .05$. The crime main effect indicated that certain crimes (e.g., rape, child molestation) were

² In a few cases (approximately 5% in each condition), jail times were either not provided or provided in non-numeric terms. These responses were removed from the analysis.

punished more severely than other crimes (e.g., credit card fraud). The main effect for crime had no theoretical significance here. The betrayal main effect was in line with theoretical predictions. The recommended jail sentences were longer in the betrayal conditions for all five offenses (see Fig. 1). These data support a betrayal effect such that people wish to punish betrayers more than non-betrayers for identical crimes.

3. Study 2: Mechanisms—two types of vulnerability and the broken promise

Having found that people wish to punish betrayers more than non-betrayers for identical harms, it is important to inquire about likely mechanisms for the phenomenon. What exactly is it about betrayals that make them more punishment-worthy?

We begin by reiterating that the betrayal harms, unlike non-betrayal harms, involve violation of a protective trust. For the form of betrayals of interest here, such violations occur when the trusted party causes the very harm he or she was supposed to guard against. Drawing on the cross-disciplinary features of trust in Rousseau et al. (1998) quoted earlier, responses to betrayal may be driven by: (a) exploitation of a vulnerable dependent party, and/or (b) violation of the vulnerable party's "positive expectations" about the behavior of the other party. Betrayals such as those considered in Study 1 typically include both an exploitation of vulnerability as well as a violation of positive expectations that was based on an implicit or explicit promise of protection. No empirical research has considered whether either the exploitation of vulnerability or the violation of a trust-based expectation serves as a psychological mechanism for betrayal responses.

3.1. Vulnerability: Access and undetectability

Vulnerability has at least two relevant components. The first is *access*: does one party have the means to exploit the vulnerability in another? To illustrate, day-care workers who commit child abuse typically have special access to their victims. This access makes children vulnerable, and exploiting this vulnerability may be the trigger for the strong punishments meted out to betrayers. If the day-care victims are extremely young, retarded, or otherwise unable to report the abuse, the crime may be difficult to detect.

The second component of vulnerability is *undetectability*: does the party who commits the bad act have reason to believe he or she is particularly likely to "get away with" the deed? Researchers in law and economics suggest that punishments should reflect the detectability of the harm that was committed (Landes & Posner, 1987; Polinsky & Shavell, 1999). Crimes that are hard to

detect require harsher punishment to compensate for the temptation to commit such crimes. However, a recent study failed to detect a relationship between probability of detection and punishment in a mock personal injury case, even when jurors' attention was directed toward this variable (Sunstein, Schkade, & Kahneman, 2000).

3.2. Violation of positive expectations: The broken promise

Aside from vulnerability due to access and/or undetectability, betrayal effects may also be caused by the violation of positive expectations that were created by a trusting relationship. Consider, for example, people who pledge (implicitly or explicitly) that they will not harm our interests or us. When such people break their promise, they do more than cause the instant harm. By violating the trust we placed in them, they threaten the very social order that permits us to have a positive expectation of safety. At the very least, their actions replace a confident sense of safety with a disturbing risk and uncertainty. This has its own consequences. Not only might it cause us to spend precious resources (e.g., energy, time, and money) developing and maintaining new safeguards,³ but it may damage our general sense in the orderliness, fairness or security of the world. This is particularly significant in light of research on the "just world hypothesis" which indicates that people have a powerful need to believe that the social world is orderly, predictable, and fair (Lerner, 1980). It may be reasonable to assign additional punishment to those who threaten the orderliness of the social world in an effort to help preserve and reinforce the norm (Durkheim, 1925/1961, p. 44; Tyler & Boeckmann, 1997; see also Tetlock, Peterson, & Lerner, 1996).

To illustrate, consider the parent who sends a child to a day-care. The parent has confident expectations about the safety of the child. These expectations derive from the implicit (and sometimes explicit) promise that the day-care will protect the child or, at the very least, not harm the child. If a day-care employee harms the child, the promise of protection that goes with entrusting one's child to the care of another is broken, and a threat to the perceived social order may be created. The parent must now spend emotional and monetary resources coping with the harm and securing more trustworthy protectors. Consequently, we offer the *broken promise* as a third potential explanation (in addition to access and undetectability) for why people may wish to punish betrayers severely.

In Study 2 we isolate and test three potential mechanisms for the desire to punish criminal betrayals

³ Evidence that people may have well evolved "cheater detection" mechanisms (Cosmides & Tooby, 1992) underscores the importance of trust and violations of trust in social exchange.

relatively severely. Does this desire stem from the protector’s exploitation of his/her special access to the protectee? Does it stem from the protector’s ability to inflict harm without detection? Does it stem from the violation of certain expectations of behavior associated with a promise not to cause harm? Study 2 tested these three potential mechanisms by systematically introducing one or more of them into a criminal scenario.

3.3. Method

3.3.1. Participants

One hundred and eighty six Columbia University students participated in Study 2. Participants were paid three dollars for their time.

3.3.2. Design and procedure

Participants were run in a classroom setting in groups ranging in size from 10 to 30. They were provided with eight randomly ordered cases of theft within the context of a 2 (scenario: groundskeeper, machine operator) × 2 (broken promise: yes, no) × 2 (access vulnerability: high, low) × 2 (indetectability: high, low) mixed design. Scenario was a between subjects variable and promise broken, access vulnerability, and indetectability were within-subject variables.

Participants were told that the eight cases were similar except for a few differences in circumstances. In the groundskeeper scenario, a man who was hired to be a groundskeeper at a store was intercepted by a police officer at 2:00 a.m. exiting the back of the store carrying several boxes of electronic equipment. The equipment was valued at \$10,000. The man was arrested, tried, and found guilty of property theft. Participants were told that theft of property valued at \$10,000 is punishable by imprisonment in the county jail for a minimum of 2 months and a maximum of 24 months. Participants were also told “case circumstances determine the exact number of months.” The machine operator scenario was similar except that the setting was a factory and the thief was a machine operator who was caught stealing equipment from a locked supply room.

The eight cases varied in terms of whether: (a) a promise was broken, (b) the store or factory was particularly vulnerable to theft by the employee (access

vulnerability), and (c) the perpetrator believed he could not get caught (indetectability). In the broken-promise condition, the (groundskeeper, machine operator) told his employer, “I promise I will be an honest and trustworthy employee.” In the no-broken-promise condition, the (groundskeeper, machine operator) told his employer, “I promise I will be an excellent (lawn cutter, machine operator).” The employee’s subsequent theft violates the promise to be trustworthy, but does not violate the promise to be an excellent (lawn cutter, machine operator).

In the groundskeeper scenario, access vulnerability was manipulated by whether the employee (a) had keys that he then used to open the store and commit the theft (high), or (b) gained access to the store by picking the lock (low). In the machine operator scenario, access vulnerability was manipulated by whether the employee (a) had a security ID that gave him access to the supply room (high), or (b) gained access to the supply room by sneaking in while others were leaving (low).

Indetectability was manipulated in the groundskeeper scenario by the employee’s knowledge about whether there would or would not be a security guard stationed at the store on the night of the theft. In the machine operator scenario, indetectability was manipulated by the employee’s knowledge about the presence or absence of security cameras in the supply room.

After reading all eight cases of theft by the groundskeeper or the machine operator, participants were provided with a second copy of the instruction sheet and bullet point summaries of the essential facts of the eight cases. Next, participants indicated how much jail time the criminal should get (in months) in each case, and explained (in writing) how they decided on the amount of punishment to assign to the different cases.

3.4. Results and discussion

A 2 (scenario: groundskeeper, machine operator) × 2 (broken promise: yes, no) × 2 (access vulnerability: yes, no) × 2 (indetectability: high, low) ANOVA on the jail time dependent measure yielded two significant effects (see Table 1). First, there was a strong main effect for the broken promise variable. Participants assigned longer jail sentences to the employee who broke his promise to

Table 1
Study 2: Mean punishment (months in jail) by proposed betrayal mechanism

Scenario	Broken promise		Access vulnerability		Indetectability	
	Yes	No	High	Low	High	Low
Groundskeeper (<i>n</i> = 98)	15.0	12.7**	13.7	14.0	14.0	13.7
Machine operator (<i>n</i> = 88)	15.7	12.1**	14.4	13.4*	13.7	14.1
Total	15.4	12.4**	14.0	13.7	13.9	13.9

* Indicates difference at *p* < .005 between the two levels of the proposed betrayal mechanism.

** Indicates difference at *p* < .001.

be an honest and trustworthy employee ($M = 15.4$ months) than to the employee who made no such promise ($M = 12.4$ months), $F(1, 184) = 76.53$, $p < .001$. No main effects for the two vulnerability variables—access vulnerability (yes: $M = 14.0$, no: $M = 13.7$) [$F(1, 184) = 1.82$, $p = .179$] and undetectability (yes: $M = 13.9$, no: $M = 13.9$) [$F(1, 184) = 0.67$, $p = .797$]—were found. However, there was an unanticipated scenario \times access vulnerability interaction [$F(1, 184) = 8.98$, $p = .003$]. Whereas an access vulnerability effect did not appear in the groundskeeper scenario (yes: $M = 13.7$, no: $M = 14.0$) [$F(1, 97) = 1.35$, $p = .248$], participants did assign longer jail sentence to the machine operator who took advantage of his easy access ($M = 14.4$ months) than to the machine operator who did not have easy access ($M = 13.4$ months), $F(1, 87) = 9.62$, $p < .005$.

One reason that we failed to detect an access vulnerability effect in the groundskeeper scenario may be that the low access vulnerability manipulation involved the commission of a second crime, namely, picking the lock. Several participants noted that lock picking is itself a criminal act that should be punished. As a result, any increase in jail time in the high access vulnerability condition may have been washed out by a countervailing effect associated with the lock picking crime in the low access vulnerability condition. The machine operator scenario manipulated access vulnerability (presence or absence of a security ID) in a way that did not involve the commission of a second crime. Here we observed an effect on jail time for access vulnerability. However, this effect was significantly *smaller* than the promise breaking effect for this scenario. Whereas thieves received an average of 1.0 additional years of jail time for exploiting access, they received 3.6 additional years for breaking a promise to be trustworthy, $t(87) = 4.44$, $p < .001$.

Participants' written explanations also pointed to the broken promise as the largest effect on recommended jail time. Two trained judges (both were graduate students at Columbia University) coded participants' written explanations into one of three primary categories: a focus on: (1) the promise (promise to be trustworthy, promise to be an excellent employee), (2) access vulnerability (used keys or ID to get in, picked lock or snuck in), and (3) knowledge about detectability (knew there would/would not be a security guard/camera present). Two additional categories that captured various other concerns (e.g., general concerns about the value of criminal incarceration) and the lack of explanations were also included. The judges also tracked the order in which the various reasons appeared for participants who mentioned more than one.

Initial agreement among judges was very good regarding the first reason participants gave to explain their judgments ($\kappa = 0.824$; Fleiss, 1981). Disagreements were resolved through consensus. Ninety-eight out of 186

participants (53%) gave reasons for their answers that could be coded in one or more of the three primary categories. Seventy of those 98 participants (71%) identified the broken promise as a first consideration. This was significantly more than either the twenty-five (26%) who identified access vulnerability ($Z = 6.43$, $p < .001$), or the three who identified knowledge about detectability ($Z = 9.90$, $p < .001$). These open-ended explanations suggest that participants were relatively more sensitive to the broken promise issue than they were to the easy access and hard-to-detect issues when deciding on punishments for the employee-thief.

In sum, Study 2 supports the idea that betrayals are punished severely, in part, because a promise to behave otherwise was broken. A theft committed by an employee who promised to behave honestly was deemed worse than one that was committed by an employee who did not make such a promise. We suspect that the groundskeeper or machine operator who failed to keep his promise is the target of harsh punishment because, like the professional protector who betrays, he creates a separate harm by defying the social norm that forbids this abuse of trust. His lie undermines the social order by introducing uncertainty and risk where there was previously a perception of safety. Study 2 also suggested that people who use their unique access to commit a crime might be punished more severely than those who do not have such access, though this effect appears to be smaller than the broken promise effect.

4. Study 3: Object betrayal

Earlier we noted that consumers form asymmetric trusting relationships with safety products such as fire alarms, vaccines, and airbags. Although safety products do not make explicit promises, we do make ourselves vulnerable to them and rely on implicit expectations that they will both protect us from and not cause the very harm they are entrusted to protect against. As in cases involving trusted human agents, these products *seem* to betray us when they cause the harms they were designed to guard against. When such apparent object betrayal occurs, we anticipate that people will respond much as they would to a human betrayal. Two expected responses are negative emotion and the desire to impose relatively strong punishments on those believed to be responsible.

These predictions are consistent with recent research that finds a strong relationship between anger and outrage and acts that deserve punishment (Bies, 1987; for a review, see Miller, 2001, pp. 534–535). For example, Kahneman, Schkade, and Sunstein (1998) argued that the severity of punishments reflects an emotional response to bad acts. They found a .78 correlation between reported outrage and median punitive damage

award in a series of hypothetical personal injury cases. In a study on anticipated emotions, Baron (1992, experiment 2) found that most people say they would feel angrier if they were blinded from a vaccine for a disease than if they were blinded from an untreated disease. Although Baron's study did not specifically invoke betrayal, the element of object betrayal that was present in the vaccine scenario may have caused the heightened emotional response.

In Study 3 we predicted that people would be more likely to assign large punitive damages against the manufacturer of a betraying safety product than they would against the manufacturer of a product that caused an identical physical harm, but which did not have an affirmative duty to protect. We also predicted that people would experience relatively more negative emotion (e.g., anger, resentment, and feelings of betrayal) toward the manufacturer of the betraying product.

4.1. Method

4.1.1. Participants

Seventy-two University of Texas students and parents who attended annual "Parents Day" activities participated in Study 3 in exchange for soft drinks.

4.1.2. Design and procedure

Participants were run individually in a classroom setting. Upon arrival, participants were randomly assigned to a betrayal or non-betrayal condition, and provided with a scenario about a fire that caused extensive damage to the structure of a warehouse. Participants were told that: (a) the fire was caused by faulty wiring in a recently installed *fire alarm* (betrayal condition) or *refrigerator* (non-betrayal condition), (b) the company responsible for manufacturing and installing the device agreed to pay for repairs and losses, and (c) the warehouse owner sued the company (fire alarm or refrigerator manufacturer) for additional money "in order to punish it." Participants were asked whether, as jurors in this case, they would "require the company to pay additional money to the business owner as additional punishment for the faulty wiring in the (fire alarm or refrigerator)." Those who answered "yes" were asked how much additional money the company should pay. Finally, all participants used a 7-point Likert-type scale to identify how much anger, resentment, and feelings of betrayal⁴ they would have if they were the plaintiff.

⁴ Although betrayal is a type of trust breach rather than a response to a breach, we asked our participants about "feelings of betrayal" to determine whether participants in the fire alarm condition were more likely than participants in the refrigerator condition to respond to the warehouse fire as a betrayal. We ask about feelings of betrayal again in Study 4.

4.2. Results and discussion

As Table 2 shows, participants in the betrayal condition (fire alarm manufacturer) meted out *more* punitive damage punishments (82% vs. 52%, $Z = 2.87$, $p < .01$) and *larger* punitive awards ($M's = \$37,200$ vs. $\$20,000$, $t(70) = 2.58$, $p < .02$) than did participants in the non-betrayal condition (refrigerator manufacturer). The size of punitive damage awards correlated with the amounts of reported "feelings of betrayal" ($r = .33$, $p < .01$), resentment ($r = .26$, $p = .03$), and anger ($r = .35$, $p < .01$).

These data support the idea that people react to apparent object betrayal with punitive punishments and negative affect. Those who imagined being harmed by a malfunctioning safety product (betrayal) were more likely to seek punitive damages than those who imagined experiencing identical harm from a malfunctioning product that did not have an affirmative duty to guard against the injury (non-betrayal). This result, and the appearance of a positive relationship between the size of damage awards and amount of feelings of betrayal, anger, and resentment, are consistent with the positive correlation between outrage and punitive damage awards reported in Kahneman et al. (1998). The data are also consistent with Buck and Miller's (1994) observation that "incongruous misfortunes" (e.g., the death of a soldier by friendly fire) generate stronger affective reactions and perceptions of injustice than equally severe congruous misfortunes (e.g., death of a soldier from enemy fire).

To this point, we have shown that: (a) people want to punish those responsible for criminal and object betrayals more severely than those responsible for identical harms that do not include a betrayal (Studies 1 and 3), (b) there is a positive relationship between punishment and negative affect (Study 3), and (c) the broken promise to protect may lie behind the desire to punish betrayers more than non-betrayers for identical immediate harms (Study 2). In Study 2, we speculated that broken promises produce betrayal responses, in part, because they undermine the social order. In Study 4, we put this speculation to the test by examining the interplay among

Table 2
Study 3: Punitive damage awards as a function of betrayal

Cause of fire	Punitive damage awards	
	Proportion who gave awards** (%)	Average size of award (\$)*
Fire alarm (betrayal) ($n = 39$)	82.1	\$37,200
Refrigerator (non-betrayal) ($n = 33$)	51.5	\$20,000

* Indicates difference at $p < .05$ between betrayal and non-betrayal conditions.

** Indicates difference at $p < .01$.

broken promises, feelings of social disorderliness, and the negative emotional reactions that people have to safety product betrayals.

5. Study 4: Broken trust, social disorder, and negative emotions

Several years ago, broad publicity was given to reports that some drivers and passengers were killed by air bags that inflated in accidents that the people otherwise would have survived (Air bag kills, 1996; Wald, 1996). Because air bags deploy with great force, the air bag itself (rather than the accident) sometimes inflicts the trauma that causes death. It is a classic case of safety product betrayal: an object that we trust to protect us actually causes the very harm (loss of life in car accidents) that it was designed to prevent. If part of the harm that results from air bag betrayals is the damage that such safety product betrayals inflict on the social order, then we should expect to see a relationship between the perception of social order damage and the experience of negative emotions.

In Study 4, we inquired about people's feelings following the (hypothetical) accidental death of a friend that either was or was not associated with a safety product betrayal. We made two primary sets of predictions. First, we predicted that people who contemplated a betrayal death compared to a non-betrayal death would be more likely to report feeling: (a) that a trust (or promise) was broken (Rousseau et al., 1998), (b) that the world is a disorderly place (Lerner, 1980), and (c) stronger negative emotions (e.g., anger, resentment, anxiety, fear, sadness, and disgust; see Lazarus, 1991). Second, we predicted that feelings of social disorder would mediate the relationship between the broken trust and the negative emotions produced by safety product betrayals.

5.1. Method

5.1.1. Participants

Undergraduate students at Columbia University were recruited through the use of advertisements placed around campus. Eighty-three students participated in the study in exchange for payment of two dollars each.

5.1.2. Design and procedure

The study was conducted in a classroom setting. Upon arrival, participants were provided with two scenarios in which an individual is killed in a car accident. In both scenarios, participants were asked to assume that they had loaned a newly purchased car equipped with an air bag to a close friend who was subsequently killed in a collision. Participants were told that an accident report revealed that the cause of death was not due to the force

of the crash itself. The cause of death was described either as "trauma caused by the force of the air bag deployment" (betrayal) or as "inhalation of toxic fumes from a damaged engine" (non-betrayal). The scenarios were presented side by side on the same page, and the order of presentation was counterbalanced.

We encouraged participants to distinguish between the two tragic deaths by asking the following forced choice question: "Obviously the death of a friend is a terrible event. Holding this issue aside, which story do you feel is worse?" Next, participants answered identical Likert-type questions (1 = strongly disagree, 7 = strongly agree) about both the first and second scenario. The order of presentation of the questions was counterbalanced with the order of presentation of the scenarios.

All questions began with the words "I would feel..." A manipulation check question asked directly about "feelings of betrayal." Three questions concerned participants' feelings about broken trust issues ("like I had been misled by something I relied on," "like my trust was violated," "like a promise had been broken"). Three questions concerned the social order ("that the world is a less orderly place," "that the world is a less certain place," "that the world is unfair"). Six questions concerned negative emotional reactions (i.e., anger, resentment, anxiety, fear, sadness, disgust; Lazarus, 1991).

5.2. Results and discussion

The manipulation check indicated that participants felt more betrayed in the betrayal (air bag deployment) condition ($M = 5.2$) than in the non-betrayal (engine fumes) condition ($M = 4.4$), $t(82) = 3.88$, $p < .001$.

5.2.1. Comparison of broken trust, social disorder, and negative emotions

We predicted that participants in the betrayal condition would report stronger feelings of broken trust, more social disorder, and stronger negative emotions than participants in the non-betrayal condition. Composite measures of feelings of broken trust, social disorder, and negative emotions were created in each condition by taking the mean of participants' responses to each of the three associated items for broken trust and social disorder and each of the six items for negative emotions. Reliabilities ranged from α s of .81 to .93, all exceeding the minimum of .70 recommended by Nunnally (1978).

Consistent with our predictions, participants in the betrayal condition reported stronger feelings of broken trust ($M = 4.3$ v. $M = 2.9$, $t(82) = 6.18$, $p < .001$), greater perceived social disorder ($M = 3.9$ v. $M = 3.7$, $t(82) = 2.15$, $p = .035$), and stronger negative emotions ($M = 4.6$ v. $M = 4.3$, $t(82) = 2.47$, $p = .015$) than participants in the non-betrayal condition (see Table 3).

Table 3
Study 4: Broken trust, social disorder, and negative emotion as a function of betrayal

Cause of death	Broken trust**	Social disorder*	Negative emotions*
Air bag (betrayal)	4.3	3.9	4.6
Engine fumes (non-betrayal)	2.9	3.1	4.3

Note. Higher scores indicate stronger beliefs and feelings ($n = 83$).

*Indicates difference at $p < .05$ between betrayal and non-betrayal conditions.

**Indicates difference at $p < .001$.

5.2.2. Mediation analysis

We predicted that the perceived social disorder that betrayals promote mediates the relationship between feelings of broken trust and the negative emotions produced by safety product betrayals. The results supported a partial mediation.

We tested the mediation hypothesis with data from the betrayal scenario (air bag deployment). According to the approach outlined in Baron and Kenny (1986), three conditions must hold in order to establish mediation: (a) the independent variable (broken trust) must affect the proposed mediator (social disorder); (b) the independent variable (broken trust) must affect the dependent variable (negative emotions); and (c) the proposed mediator (social disorder) must affect the dependent variable (negative emotions). If all three conditions hold, then the effect of the independent variable (broken trust) on the dependent variable (negative emotions) must be reduced when the proposed mediator (social disorder) is included in the regression equation. Our results suggest that feelings of social disorder mediate the relationship between broken trust beliefs and negative emotions (see Fig. 2).

First, a regression of participants' feelings of broken trust as the independent variable and negative emotions as the dependent variable revealed a significant relationship, $t(82) = 4.42$, $p < .001$. Second, feelings of broken trust were also related to feelings of social disorder, $t(82) = 3.74$, $p < .001$. Finally, including both

broken trust and social disorder as independent variables in a regression predicting negative emotions revealed a significant beta for disrupted social order, $t(82) = 6.54$, $p < .001$, as well as for broken trust, $t(82) = 2.53$, $p = .014$. Because the standardized beta for betrayal ($\beta_{\text{broken trust}} = .436$) decreased with the inclusion of social disorder in the model ($\beta_{\text{broken trust}} = .220$), the data supported a partial mediation. The mediation was significant, $Z = 3.25$, $p < .001$ (Kenny et al., 1998; Sobel, 1982).

In sum, Study 4 reinforced and extended the results of our previous studies related to the psychology of betrayal. First, it showed that betrayal by a safety device is associated with a greater sense of broken trust, more perceived social disorder, and stronger negative emotions than an identical negative outcome that was not caused by a betrayal. These object betrayal results are consistent with Study 2 which found that the element of broken trust in criminal betrayal led to greater recommended punishments. They are also consistent with Study 3 which found that a harm that includes a betrayal was associated with greater negative emotions than an identical non-betrayal harm. Second, mediation analyses supported our hypothesis that the negative emotions associated with betrayal emerge, in part, from the perception of social disorder that the broken trust of a betrayal creates. Thus, one could argue that betrayals evoke strong negative reactions, in part, because they harm us in multiple ways. Not only do they cause the focal harm associated with the offense, but they also undermine our sense that the social world is fair and orderly (Lerner, 1980). This conclusion fits well with recent research that traces the public's punitive feelings toward repeat offenders to a belief that such people are an affront to the "moral cohesion" and "social consensus" of society (Tyler & Boeckmann, 1997; see also Darley, Carlsmith, & Robinson, 2000).

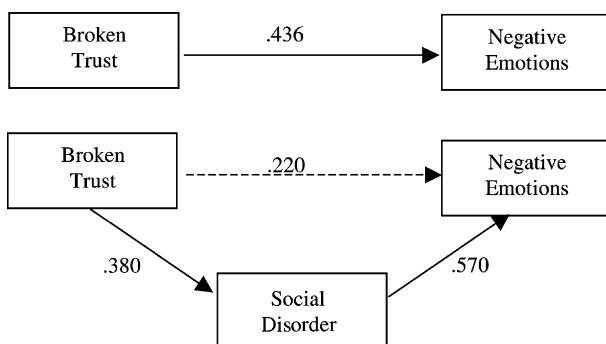


Fig. 2. Study 4: Social disorder as a mediator of broken trust and negative emotion. Note. The path diagrams show that social disorder partially mediates the association between broken trust and negative emotions. Path coefficients shown are standardized regression coefficients. All paths are $p < .001$.

6. Study 5: Aversion to safety product betrayal risks

Thus far, our focus has been on people's response to betrayals *after* they have occurred. In a final study, we consider people's response to the risk of future betrayals. If the real and perceived costs of betrayals are sufficiently large, people might be willing to incur serious

costs to prevent betrayals from arising in the first place. For example, employers concerned about employee betrayals might monitor office activities with expensive surveillance devices. In Study 5, we extend this idea to the consumer domain by considering the lengths that people will go to eliminate the risk of a safety product betrayal. Specifically, we investigate whether people would be willing to accept an increased risk of the very harm they seek protection from to eliminate the mere *possibility* of a safety product betrayal.

For example, suppose that a consumer must choose between two equally-priced safety products. If the chance of harm associated with product A is twice that associated with product B (e.g., 2% vs. 1%), consumers would obviously select product B. But what if product B is also associated with a very small risk of betrayal (e.g., 0.01%)? Although a probabilistic approach favors option B (2% vs. 1.01%), betrayal aversion may lead consumers to choose the riskier product A. That is, aversion to small betrayal risks in safety products may be so great that consumers avoid these options in favor of options that put them at *greater* chance for harm.

This hypothesis is consistent with a growing body of literature that shows how various psychological considerations affect risky decision-making. It is well-established that risk judgments are informed by a risk's voluntariness, controllability, familiarity, visibility, future threat, catastrophic potential, reversibility, dreadfulness, outrage, and frame (Slovic, 1987, 1997; Slovic, Fischhoff, & Lichtenstein, 1977, 1982; Vlek & Stallen, 1981; Wandersman & Hallman, 1993; Zeckhauser & Viscusi, 1990). Some of these variables may trigger affective reactions that influence whether people are willing to expose themselves to risks (see e.g., Isen & Patrick, 1983). Indeed, recent research suggests that affective reactions play a central role in decision making (Finucane, Alhakami, Slovic, & Johnson, 2000; Luce, 1998; Mellers, Schwartz, Ho, & Ritov, 1997; Mellers, Schwartz, & Ritov, 1999; Slovic, 2000).⁵ For example, Loewenstein, Weber, Hsee, and Welch (2001) have argued that emotional or gut level responses to risks can have a larger impact on judgment and choice than cold cognitive evaluations. In support, they noted that emotional reactions (including, presumably, reactions to betrayals) are sensitive to the mere *possibility* of negative consequences (for a similar point about risk perception more generally, see Teigen, Brun, & Frydenlund, 1999).

⁵ Whether affect *should* play a central role in decision-making is a more dubious proposition. Recent social psychological research shows that people overestimate the duration of their affective reactions to a variety of negative outcomes (Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998) and underestimate the impact of outcome surprise on how good or bad we will feel (Coughlan & Connolly, 2001). If true, then the affective cues that people use to inform risky choice may be too heavily weighted against those options that are associated with the most negative outcomes.

Reviewing evidence from early psychophysical studies, Loewenstein et al. (2001) pointed out that “the mere thought of receiving a shock was enough to arouse participants, and the precise likelihood of being shocked had little impact on their arousal level” (p. 276; see also Damasio, 1994). By extension, this suggests that even when the probability of a betrayal is quite small (e.g., 0.01%), we might expect that people will react to and be influenced by its mere presence.

Research on protected values and taboo tradeoffs also shows that options that include a risk of betrayal may be rejected by an affect-based system of reasoning. Baron and Spranca (1997); (see also Ritov & Baron, 1999) argued that certain noncommodity attributes (e.g., health, honesty, and life) are “protected” in the sense that we are reluctant to trade them off with other values. Indeed, people often become angry when asked to exchange a protected value for money. Fiske and Tetlock (1997) argued that such exchanges are viewed as so morally degrading and threatening to the social order that they are taboo. Similarly, Slovic and his colleagues have suggested that certain risks and actions are stigmatized as unacceptable at any price (Kunreuther & Slovic, 1999; Slovic, Flynn, & Gregory, 1994).

In sum, an emerging literature suggests that affect, moral beliefs, and other non-rationalist considerations influence how people think about risky decision making including, presumably, betrayal risks. This literature provides theoretical grounding for our prediction that the mere possibility of betrayal in a safety device may be so aversive that people will avoid it in favor of an option that affords less protection.

Recent events suggest this may be true in the natural ecology. As noted earlier, there is some evidence that air bags sometimes kill people in otherwise survivable accidents. However, the evidence also shows that nearly all drivers (including short ones) are safer with a driver's side air bag than without one (National Highway Traffic Safety Administration, 1999; Wald, 1997).⁶ Nevertheless public anxiety about air bag dangers has soared, leading the US government to order auto manufacturers to install toggle switches in vehicles that would allow motorists to shut off air bags (Ball, 1999). Thousands of drivers requested the toggle switches when they became available (Air bag on or off?, 1998; Switching off the air bag, 1997).

In Study 5, we examine betrayal risk aversion in the context of hypothetical choice scenarios involving three safety products: air bags, smoke alarms, and vaccines.

⁶ Only the small proportion (2.5%) of people who do not ordinarily achieve a 10 inch distance between their breastbone and the air bag when driving cover are at risk from air bags. Even within this group, an sample of short women (4'8"–5'2") were able to adjust their driving position to achieve the 10 inch distance in nearly all test vehicles (National Highway Traffic Safety Administration, 1999).

The air bag scenario is similar to the one used in Study 4 where the potential for death was represented by an exploding air bag (betrayal) or fumes from a damaged engine (non-betrayal). The smoke alarm scenario was similar to the one used in Study 3 where the betrayal harm was represented by a fire alarm that caused a fire. The vaccine scenario was loosely based on stimuli used in Ritov and Baron (1990) and Asch et al. (1994).

6.1. Method

6.1.1. Participants

Four hundred and eighty-one undergraduate students at the University of Texas at Austin participated in the study in exchange for partial course credit in introductory business classes.

6.1.2. Design and procedure

Participants were run in groups ranging in size from 20 to 80 in a classroom setting. At the beginning of the class, participants were told that we were conducting a decision making study in which there were no right or wrong answers.

We used a fully crossed 3 (scenario: air bag, vaccine, smoke alarm) \times 2 (betrayal: yes, no) between-subjects design. Three hundred and eighty-three participants were assigned at random to one of the six experimental conditions ($n = 62$ – 66 per condition). In addition, 98 participants were randomly assigned to one of three control conditions: air bag ($n = 33$), vaccine ($n = 34$), smoke alarm ($n = 31$).

Participants in the six experimental conditions were presented with a choice between a single risk and a dual risk safety option using stimuli that were very similar to those used in Study 4. In this case, Brand A was associated with a 2% risk of death and Brand B was associated with a 1% risk of death plus an additional 1/10,000 (0.01%) risk of death. The additional 0.01% risk in Brand B was described either as a betrayal risk (i.e., the chance that Brand B itself might cause death) or a non-betrayal risk (i.e., the chance that death might occur in association with Brand B, but not because of it). To illustrate, participants in the *air bag betrayal* condition read the following passage:

Suppose that you are offered a choice between two equally priced cars: Car A and Car B. Car A is equipped with Air Bag A. Scientific crash tests indicate that there is a 2% chance that drivers of Car A who are in serious accidents will be killed due to the impact of the crash. Car B is equipped with Air Bag B. Scientific crash tests indicate that there is a 1% chance that drivers of Car B who are in serious accidents will die due to the impact of the crash. However, Car B may kill drivers who would not have died if they were driving Car A instead. Specifically,

some drivers of Car B may die due to trauma caused by the force of the air bag deployment. Crash tests indicate that there is an additional one chance in 10,000 (0.01%) that someone who is in a serious accident in Car B will be killed due to air bag trauma.

In the *air bag non-betrayal* condition, the risk of death due to air bag deployment was replaced by a risk of death due to toxic fume inhalation from a damaged engine. Unlike air bag deployment trauma, toxic fume inhalation is not a betrayal of the safety device itself. Key betrayal and non-betrayal manipulations from each of the three scenarios are presented in the Appendix A.⁷

Participants in all six experimental conditions and all three control groups were asked: (a) which brand they would prefer (forced choice), (b) how strong their brand preference was (using a 7 point Likert-type scale in which low values favor the high risk option), and (c) to explain their answers in writing. We predicted a main effect for betrayal on all dependent measures in all three scenarios such that people would favor the low risk option (Brand B) in the non-betrayal condition, but favor the high risk option (Brand A) in the betrayal condition.

Participants in the three control conditions (one for each scenario) were not asked to choose between a single risk (2%) option and a dual risk (1+0.01%) option. Instead, they were asked to choose between two single risk options: a 2% death risk vs. a 1.01% death risk. The control conditions served as a check that participants were sensitive to risk differences at these probability levels. They also provided a baseline from which to compare responses in the betrayal and non-betrayal conditions within each scenario.

We predicted that people would be willing to accept increased risks of a negative outcome to avoid a very small risk that the identical outcome arises from a betraying safety product. We also predicted that this effect would disappear when the small betrayal risk was replaced by an outcome-identical non-betrayal risk.

6.2. Results and discussion

Predictions were tested using a binary logistic model. As predicted, there was a main effect for betrayal in

⁷ Slight variations were introduced in the three scenarios to enhance the generality of the experiments. In the air bag and vaccine scenarios, participants were asked to imagine that a potentially life-threatening event (i.e., a car accident and a disease) may occur in which a safety product is enlisted for aid. In the smoke alarm scenario, participants were asked to imagine selecting a safety device prior to the appearance of the potentially life-threatening condition. In all three scenarios, the betrayal risk captured the risk of death caused by a safety product in situations that would have turned out favorably if the safety product were absent.

participants' safety device choices (Wald $\chi^2(1) = 7.63$, $p < .01$). Table 4 shows that only 32.6% of participants in the betrayal condition preferred Brand B (1% death risk + 0.01% betrayal death risk) to Brand A (2% death risk). It appears that most people are willing to double their overall chance of dying to avoid incurring a very small chance of dying via betrayal. But when the 0.01% additional risk associated with Brand B was described as a non-betrayal risk, a different pattern emerged. Most participants in this non-betrayal condition (61.7%) preferred Brand B to Brand A. In other words, when the betrayal component of risk was replaced by a statistically identical non-betrayal component, most participants chose the safety device that minimized their overall risk of dying. This pattern of results was consistent across all three scenarios (see Table 4).

A 2 (betrayal: yes, no) \times 3 (scenario: air bag, vaccine, smoke alarm) ANOVA was used to examine differences in participants' strength of preferences for safety devices across conditions. Participants' strength of preferences mirrored the choice results. In the betrayal condition, participants' mean strength of preference tended toward Brand A ($M = 3.4$). In the non-betrayal condition, strength of preference tended toward Brand B ($M = 4.4$). This difference was significant, $F(1, 377) = 30.98$, $p < .001$. This pattern of results was consistent across all three scenarios (see Table 5).

Participants' open-ended explanations of their choices provided additional support for betrayal aversion.

Table 4
Study 5: Proportion who chose Brand B (1% +0.01% risk of death) over Brand A (2% risk of death)

Scenario	Betrayal	Non-betrayal
Air bag	27.4% (17/62)***	63.6% (42/66)
Alarm	25.0% (16/64)**	51.6% (33/64)
Vaccine	45.3% (29/64)*	69.8% (44/63)
Total	32.6% (62/190)*	61.7% (119/193)

* Indicates difference at $p < .01$ from non-betrayal condition.

** Indicates difference at $p < .005$.

*** Indicates difference at $p < .001$.

Table 5
Study 5: Mean strength of preference for Brand B (1 + 0.01% risk of death) over Brand A (2% risk of death)

Scenario	Betrayal	Non-betrayal
Air bag	3.19***	4.67
Alarm	3.00**	3.96
Vaccine	4.00*	4.66
Total	3.40***	4.43

Note. Higher scores indicate a stronger preference for the lower risk safety device (Brand B).

* Indicates difference at $p = .057$ from non-betrayal condition.

** Indicates difference at $p < .005$.

*** Indicates difference at $p < .001$.

Two trained judges coded participants' explanations into one of three categories: concern about the additional (0.01%) risk, concern about overall risk, and concern about other matters. Initial agreement among judges was excellent ($\kappa = 0.76$; Fleiss, 1981). Disagreements were resolved by a third judge and through consensus. The results showed that participants in the betrayal conditions were more likely to express concern about the additional 0.01% death risk than participants in the non-betrayal conditions. Fifty-eight percent of participants (111/190) in the betrayal conditions expressed concern about the additional betrayal risk. Typical explanations among those in the betrayal conditions were "If [the air bag] is supposed to help me stay alive, then I want the one that isn't going to kill me on it's own," and "I'd hate to have the air bag that's supposed to save my life be the cause of its termination." In contrast, only 32% of participants (62/193) in the non-betrayal conditions expressed concern about the additional non-betrayal risk, $Z = 5.36$, $p < .001$. Most participants here focused on the probability of death differential between the options. Typical explanations among those in the non-betrayal conditions were "Because with Car B there is a lower chance (1.01%) that drivers who are in accidents will die" and "Why not improve your odds [of living] while you easily can?"

Nearly all participants in the control conditions preferred the low risk alternative (93/98 = 95%) and preferred it strongly ($M = 6.2$). This indicates that our participants were not indifferent between a 2% and a 1.01% risk of death. Absent other considerations, participants had a strong preference for lower risk options at these probability levels. We note that degree of preference for Brand B is stronger in the control group than in the non-betrayal condition (choice: 95% vs. 62%, $Z = 6.03$, $p < .001$; strength of preference, $M = 6.2$ vs. $M = 4.4$, $t(289) = 8.76$, $p < .001$). This suggests that betrayal was not the only factor that influenced choices. Several possibilities, including a preference for fewer enumerated risks, are identified below.

In sum, the results of Study 5 provide strong support for betrayal aversion in choice among safety devices. Most participants (67.4%) preferred the high risk safety option (Brand A: 2% risk) when the low risk option (Brand B: 1 + 0.01%) carried with it a small additional risk of betrayal. But when the same small additional risk was not associated with betrayal, most participants (61.7%) selected the low risk safety option. These data suggest that betrayal risks are treated differently from non-betrayal risks. Indeed, betrayal risks appear to be so psychologically intolerable that people are willing to double their risk of death from automobile crashes, fires, and diseases to avoid a small possibility of death by safety device betrayal.

6.2.1. Alternative explanations

At this point, we consider five alternative explanations for the betrayal aversion detected in Study 5. These include preference for fewer risks, omission bias, regret, counterfactual reasoning, and controllability. We conclude that, while some may exert an influence over choice, they are insufficient to explain the pattern of results observed here.

6.2.1.1. Preference for fewer risks. Because the betrayal option was presented as two risks (1 + 0.01%) and the non-betrayal option was presented as a single risk (2%), it is tempting to suggest that our data are a byproduct of a risk-counting heuristic (i.e., “two risks are worse than one risk, so I’ll choose the one risk option” (cf. Josephs, Giesler, & Silvera, 1994)). However, most participants selected the dual risk option when the 0.01% risk appeared as a non-betrayal risk. It may be that the greater preference for the low risk option among participants in the control condition versus those in the non-betrayal condition reflects a preference for safety options that have fewer enumerated risks. However, this explanation does not account for the observed differences between the betrayal and non-betrayal conditions.

6.2.1.2. Omission bias. Betrayal aversion is different from omission bias. Omission bias predicts that bad outcomes that arise from actions are worse than bad outcomes that arise from inactions (Ritov & Baron, 1999; see also Feldman, Miyamoto, & Loftus, 1999). Betrayal aversion predicts that bad outcomes that arise from exploitation of protective trusts will be considered worse than identically bad outcomes that arise from actions that do not involve exploitation of a protected trust. We observed betrayal aversion when participants actively selected between two safety devices (as opposed to, say, selecting between a safety device and a status quo in which there was no safety device). It is not clear that omission bias makes predictions about choice when both choices alter the status quo.

6.2.1.3. Regret and anticipated regret. Much research shows that people prefer options that minimize the possibility of regret (Bell, 1982; Gilovich & Medvec, 1995; Loomes & Sugden, 1982; Simonson, 1992; Zeelenberg et al., 1998). Although there is disagreement about what regret entails (Connolly, Ordóñez, & Coughlan, 1997; Ordóñez & Connolly, 2000), it is often described as a negative emotion that is experienced when we realize or imagine that an option not chosen would have produced a better outcome (Zeelenberg, 1999). Research shows that people make choices that minimize the risk of regret (Zeelenberg, Beattie, van der Plicht, & de Vries, 1996) and that they are willing to pay to avoid it (Bell, 1983; Larrick & Boles, 1995). However, this research does not explain why outcomes associated with

betrayals are considered worse than the *identical outcomes* that arise in a different way. In Study 4, we showed that betrayal gives rise to various negative emotions,⁸ and this process is partially mediated by a perception that the betrayals disrupt the social order. So while anticipated regret predicts that people will be averse to certain risky options, it does not by itself explain why the risky option associated with a betrayal is avoided more than an equally risky option that is associated with a non-betrayal.

6.2.1.4. Counterfactual reasoning. Research on counterfactual reasoning by Miller and his colleagues suggests that outcomes arising from unusual events elicit more extreme reactions than those that arise in more usual ways (Buck & Miller, 1994; Kahneman & Miller, 1986; Miller & Turnbull, 1990). The betrayal options in Study 5 were associated with the possibility of surprising or unusual outcomes (i.e., death by air bag trauma, fire, or vaccine-induced disease). Therefore, these options may have elicited a heightened reaction which, in turn, made them exceedingly unattractive. However, the non-betrayal options were also associated with surprising or unusual outcomes (i.e., death by engine fumes, a collapsed ceiling, and infection). Therefore it is not clear that counterfactual reasoning accounts for our results.

6.2.1.5. Controllability. It is well known that people are more willing to take risks when they believe they have some control over those risks than when they do not (Langer, 1975). The scenarios and options presented in Study 5 varied in degree of controllability. For example, some participants may have felt that the car accident scenario was controllable and that the probabilities provided were not relevant to their choices. However, the vaccine scenario is not a particularly controllable one because participants had no prior knowledge about the fictional disease. Because the pattern of results did not vary across scenarios, controllability does not provide a compelling account for our data.

In sum, the alternative explanations identified above probably do not explain the pattern of betrayal aversion that we observed across the three safety product scenarios in Study 5.

7. Summary and general discussion

At the outset, we noted that researchers have yet to agree on a simple definition of betrayal. In this paper, we focused on how people respond to the subset of

⁸ Although we did not measure it, regret (or anticipated regret) may be another negative emotion that people experience following a betrayal.

betrayals that occur when the objects of our trust cause the very harm they were entrusted to guard against. The security guard who steals and the automobile air bag that kills were identified as examples of such betrayal by a person and by a safety product respectively. Based on the participants' responses to criminal betrayals, safety device betrayals, and the risk of future betrayals in five studies that employed various experimental designs, betrayal aversion appears to be a robust phenomenon.

Study 1 found that acts of criminal betrayal arouse a stronger desire to punish than the identical criminal acts that do not involve betrayal in a mixed design experiment with betrayal manipulated between subjects. Study 2 tested three possible mechanisms (broken promise, access vulnerability, undetectability) for the betrayal punishment response in two theft scenarios in a mixed design experiment with the elements of betrayal manipulated within subjects. A strong effect for broken promise on desired punishment was found, where the broken promise was operationalized in a scenario in which employees robbed their employers shortly after promising to be honest and trustworthy. We suggested that one reason broken promises are punished severely is that they undermine the fragile trust and order upon which our social world is built. Study 3 demonstrated that people respond to object betrayals (i.e., safety products that caused the very harm they were designed to protect against) much as they do to human betrayals. A between-subjects experiment showed that people assigned larger punitive damage awards and reported stronger negative feelings against the manufacturer of a betraying product than against the manufacturer of a non-betraying product that caused identical harm. Having established an association between punishment and negative affect in an object betrayal setting, Study 4 produced evidence that the broken trust and negative emotions that accompany a safety product betrayal are (partially) mediated by the perceived breakdown in social order that betrayals promote. In Study 5, we examined the possibility that betrayals are so aversive that people may be willing to accept an increased risk of the very harm they wish to avoid to eliminate the mere possibility of experiencing that harm via betrayal. We showed that when faced with a choice among pairs of safety products (i.e., air bags, smoke alarms, or vaccines), most people selected inferior products (in terms of overall risk exposure) over those that were associated with a slim chance of betrayal. However, when the slim betrayal risk was replaced by an equivalent non-betrayal risk, the choice pattern was reversed, and a majority of people preferred the safety product that was associated with the smallest overall risk exposure. This result extended the findings in Study 4 which, like the first scenario in Study 5, compared responses to death by air bag (betrayal) and engine fumes (non-betrayal) in an automobile accident. If people feel betrayed in cases

of air bag accidental death, and if betrayal is associated with various negative emotions and concerns about social disorder, it follows that people will be willing to incur a cost to eliminate the betrayal risk. That people are willing to accept an increased risk of the very thing they wish to prevent (death) to eliminate the mere chance of betrayal strikingly illustrates the depth of betrayal aversion.

The five studies reported here approached the betrayal aversion phenomenon using various scenarios, experimental designs, and dependent measures. Taken together, they provide the first empirical evidence for people's special aversion to harm that is caused by the very agent that is charged with protecting them from such harm.

Although multiple factors may be at work, the desire to avoid such potential harms (which we have referred to as "betrayals") and to punish the responsible parties when they occur partly reflect the importance of trust and order in the social world. If our positive expectation that agents of protection will do no harm is undermined, risk and uncertainty appear where we once believed protection was assured. We draw additional support for this claim from research in decision theory and the law that shows that people are reluctant to accept even minimal risks when the alternative is zero risk. For example, the "certainty effect" shows that people are willing to pay relatively large insurance premiums to eliminate very small risks of harm (Tversky & Kahneman, 1981). Similarly, legal decision makers assign their highest monetary damage awards in medical negligence cases when an apparently riskless health situation (i.e., 0% risk of death) is turned into a risky one (e.g., 20% risk of death; Koehler, in press). We therefore suggest that the disruption to stable, certain aspects of our social world that betrayals promote lies at the core of our desire to punish betrayals when they occur and to eliminate betrayal risks where possible.

7.1. The normative issue

Many have argued that psychological and emotional influences on risk perception and evaluation are not necessarily indicative of poor thinking (Evans, 1994; Ritov, 1994; Slovic, 1987; Tetlock, 1994). Likewise, betrayal aversion may be defensible, particularly if one considers consequences that reach beyond the focal scenario. Because trust lost is not easily restored (Forward & Bruck, 1978; Rothbart & Park, 1986), the negative consequences of betrayals may be protracted and severe. This may justify the betrayal effects we found. On the other hand, care must be taken not to allow this generic argument to interfere with people's ability to make choices that maximize consistency with their goals (Baron, 1994).

7.2. Future research

Future research on betrayal aversion may wish to consider moderators of the effects presented here. For example, the consequences of betrayal and the breadth of the trust that was violated are two potentially important variables. Regarding betrayal consequences, the present research focused on high-consequence betrayals (e.g., automobile accidents). Other large stake betrayals might include those that involve risks to our health, happiness or careers. Smaller stakes betrayals might include headache medicines that sometimes cause headaches (Parker-Pope, 2000) or conversational betrayals in which a trusted friend says something hurtful.⁹ Our will to avoid and/or punish low stakes betrayals is an open empirical question.

Regarding the breadth of trust that was violated, it may be that safety products that violate a narrowly construed duty to protect are more likely to be treated as betrayals than products that violate a broader protection mandate. Consider, for example, a drug taken to reduce the risk of breast cancer (Tamoxifen) that increases the user's risk of uterine cancer (Pear, 1998). For women who construe the drug's purpose narrowly, the development of uterine cancer following Tamoxifen treatment may not seem like a betrayal. They might only feel betrayed if the drug increased the risk of breast cancer. However, for women who construe the protective purpose of Tamoxifen more broadly (e.g., "Tamoxifen protects against cancer"), the development of uterine cancer following Tamoxifen treatment may seem like a betrayal.

Additional research is needed to examine factors that affect the breadth of people's construals of protective trusts. Such research may have implications for legal judgments of liability in cases where the violation of a trust caused harm. To the extent that jurors construe the protective purpose of protective trusts broadly, they may be more likely to see betrayal and to punish offenders severely.

Finally, future researchers may wish to examine how people respond to betrayals other than those in which agents of protection caused the very harm they were entrusted to guard against. If, as some have suggested (Elangovan & Shapiro, 1998; Morris & Moberg, 1994), betrayal occurs whenever a trustee violates key expectations of a trustor in ways that may harm the trustor, then betrayal aversion may be at least partly responsible

for a variety of behaviors that appear to violate normative theory.

Acknowledgments

The authors thank Jon Baron, Rachel Crosen, Liz Thompson Gershoff, Brian Gibbs, Daniel Gilbert, Eric Johnson, Barb Mellers, Molly Mercer, Phil Tetlock, Kathleen Valley, Elke Weber, Patricia West, and Jim Westphal for helpful comments along the way. We also thank Vildan Altuglu, Sejal Haribhakti, Raghuram Iyengar, Anirban Mukhopadhyay, and Caryn Thesing for superb research assistance. Portions of this paper were presented at the ARG workshop at Harvard Business School, the Columbia University Social Psychology Network, the Columbia Business School Marketing Camp, the Decision Sciences Workshop in the Wharton School of Business, the Quantitative Workshop at Ohio State University, and the Social Psychology Research Workshop at Harvard University.

Appendix A. Key sentences from three scenarios in study 5

A.1. Air bag scenario

However, Car B may kill drivers who would not have died if they were driving Car A instead. Specifically, some drivers of Car B may die due to [trauma caused by the force of the air bag deployment (*betrayal*)/inhalation of toxic fumes from a damaged engine (*non-betrayal*)]. Crash tests indicate that there is an additional one chance in 10,000 (i.e., 0.01%) that someone who is in a serious accident in Car B will be killed by [air bag trauma/toxic fume inhalation].

A.2. Smoke alarm scenario

However, Apartment B may lead to the death of people who would not have died if they were living in Apartment A instead. Specifically, some people in Apartment B may be killed due to [an electrical fire caused by the smoke alarm wiring (*betrayal*)/ the impact caused by a collapsed ceiling (*non-betrayal*)]. Studies indicate that there is an additional one chance in 10,000 (i.e., 0.01% chance) that someone will be killed due to [an electrical fire caused by the smoke alarm wiring/the impact caused by a collapsed ceiling].

A.3. Vaccine scenario

However, some people who are treated with Vaccine B, and who would not have died if they were treated with Vaccine A, may die in another way. Specifically,

⁹ The interpretive element of social betrayals also makes this an interesting area for further investigation. One friend may think another has betrayed her when she answers a question honestly instead of supportively. But the answering friend may not see this as betrayal and may even believe a failure to answer honestly would have been a betrayal.

some people who are treated with Vaccine B may die due to [complications induced by the vaccine (*betrayal*)/infection (*non-betrayal*)]. Medical tests indicate that there is an additional one chance in 10,000 (i.e., 0.01%) that someone who has [the hypothetical disease] and who is treated with Vaccine B will die due to [vaccine-induced complications/infection].

References

- Air bag kills child secured by a seat belt (1996, October 24). Associated Press.
- Air bag on or off? Now some can choose (1998, January 19). CNN.
- Akerstrom, M. (1991). *Betrayal and betrayers: The sociology of treachery*. New Brunswick, NJ: Transaction Publishers.
- Allen, T. (1999). Muzzled by the CIA. *Legal Times*, 54.
- Asch, D. A., Baron, J., Hershey, J. C., Kunreuther, H., Meszaros, J., Ritov, I., & Spranca, M. (1994). Omission bias and pertussis vaccination. *Medical Decision Making*, 14, 118–123.
- Asmus, M. J., Barros, M. D., Liang, J., Chesrown, S. E., & Hendeles, L. (2001). Pulmonary function response to EDTA, an additive in nebulized bronchodilators. *Journal of Allergy & Clinical Immunology*, 107, 68–72.
- Baker, T. (1994). Constructing the insurance relationship: Sales stories, claims stories, and insurance contract damages. *Texas Law Review*, 72, 1395–1433.
- Ball, J. (1999). High-tech air bags are lacking in grasp of human dimensions. *Wall Street Journal*, A1, A6.
- Barber, B. (1983). *The logic and limits of trust*. New Brunswick: Rutgers University Press.
- Barnum, (1998). Home-filtered water fails lead test: System brands named in S.F. study cover half the market. *San Francisco Chronicle*, A1.
- Baron, J. (1992). The effect of normative beliefs on anticipated emotions. *Journal of Personality and Social Psychology*, 63, 320–330.
- Baron, J. (1994). Nonconsequentialist decisions. *Behavioral and Brain Sciences*, 17, 1–42.
- Baron, J., & Spranca, M. (1997). Protected values. *Organizational Behavior and Human Decision Processes*, 70, 1–16.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Bell, D. (1982). Regret in decision making under uncertainty. *Opinion Research*, 20, 961–981.
- Bell, D. (1983). Risk premiums for decision regret. *Management Science*, 29, 1156–1166.
- Bies, R. J. (1987). The predicament of injustice: The management of moral outrage. *Research in Organizational Behavior*, 9, 289–319.
- Buck, M. L., & Miller, D. T. (1994). Reactions to incongruous negative life events. *Social Justice Research*, 7, 29–46.
- Buunk, B. (1982). Strategies of jealousy: Styles of coping with extramarital involvement of the spouse. *Family Relations*, 31, 13–18.
- Connolly, T., Ordonez, L. D., & Coughlan, R. (1997). Regret and responsibility in the evaluation of decision outcomes. *Organizational Behavior and Human Decision Processes*, 70, 73–85.
- Cosmides, L., & Tooby, J. (1992). In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind*. Oxford: Oxford University Press.
- Coughlan, R., & Connolly, T. (2001). Predicting affective responses to unexpected outcomes. *Organizational Behavior and Human Decision Processes*, 85, 211–225.
- Darley, J. M., Carlsmith, K. M., & Robinson, P. H. (2000). Incapacitation and just deserts as motives for punishment. *Law and Human Behavior*, 24, 659–683.
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason and the human brain*. New York: Avon.
- Davis, J. L., & Petretic-Jackson, P. A. (2000). The impact of sexual abuse on adult interpersonal functioning: A review and synthesis of the empirical literature. *Aggression and Violent Behavior*, 5, 291–328.
- Durkheim, E. (1925/1961). *Moral education: A study in the theory and application of the sociology of education*. New York: Free Press.
- Elangovan, A. R., & Shapiro, D. L. (1998). Betrayal of trust in organizations. *The Academy of Management Review*, 23, 547–566.
- Evans, J. B. T. (1994). Normative and descriptive consequentialism. *Behavioral and Brain Sciences*, 17, 15–16.
- Feldman, J., Miyamoto, J., & Loftus, E. F. (1999). Are actions regretted more than inactions. *Organizational Behavioral and Human Decision Processes*, 78, 232–255.
- Finkelhor, D., & Browne, A. (1985). The traumatic impact of child sexual abuse: A conceptualization. *American Journal of Orthopsychiatry*, 55, 530–541.
- Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, 13, 1–17.
- Fiske, A., & Tetlock, P. E. (1997). Taboo trade-offs: Reactions to transactions that transgress spheres of justice. *Political Psychology*, 18, 255,297.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed). New York: Wiley.
- Forward, S., & Bruck, C. (1978). *Betrayal of innocence*. New York: Penguin.
- Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, 75, 617–638.
- Giles, S. M. (1995). Promises betrayed: Breach of confidence as a remedy for invasions of privacy. *Buffalo Law Review*, 43, 1–101.
- Gilovich, T., & Medvec, V. H. (1995). The experience of regret: What, when and why? *Psychological Review*, 102, 379–395.
- Isen, A. M., & Patrick, R. E. (1983). The effect of positive feelings on risk-taking: When the chips are down. *Organizational Behavioral and Human Decision Processes*, 31, 194–202.
- Josephs, R. A., Giesler, R. B., & Silvera, D. H. (1994). Judgment by quantity. *Journal of Experimental Psychology: General*, 123, 21–32.
- Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, 92, 136–153.
- Kahneman, D., Schkade, D., & Sunstein, C. R. (1998). Shared outrage and erratic awards: The psychology of punitive damages. *Journal of Risk and Uncertainty*, 16, 49–86.
- Kenny D. A., Kashy D. A., & Bolger, N. (1998). Data analysis in social psychology. *The handbook of social psychology*. Boston: McGraw-Hill.
- Koehler, J. J. (in press). Which chance was lost? The psychology of damage awards under the loss of chance doctrine. To appear in: I. Brocas & J. Carillo (Eds.), *Psychology and Economics: Selected works*, Vol. 2. Oxford: Oxford University Press.
- Kunreuther, H., & Slovic, P. (1999). Coping with stigma: challenges and opportunities. *Risk: Health, Safety and Environment*, 10, 269–280.
- Landes, W. M., & Posner, R. A. (1987). *The economic structure of tort law*. Cambridge, MA: Harvard University Press.
- Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, 32, 311–328.
- Larrick, R. P., & Boles, T. L. (1995). Avoiding regret in decisions with feedback: A negotiation example. *Organizational Behavior and Human Decision Processes*, 63, 87–97.

- Larson, L. R. (1993). Betrayal and repetition: Understanding aggression in sexually abused girls. *Clinical Social Work Journal*, 21, 137–150.
- Lazarus, R. S. (1991). *Emotion and adaptation*. New York: Oxford University Press.
- Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. New York: Plenum Press.
- Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, E. S. (2001). Risk as feelings. *Psychological Bulletin*, 127, 267–286.
- Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, 92, 805–824.
- Luce, M. F. (1998). Choosing to avoid: Coping with negatively emotion-laden consumer decisions. *Journal of Consumer Research*, 24, 409–433.
- Matthews, M. (1999). A snarl over bear repellent leaves consumers confused. *The Washington Post*, 1, A9.
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20, 709–734.
- Medvec, V. H., Valley, K. L., Thaler, R. (1999). Concession aversion: A story of loss and betrayal. Unpublished manuscript.
- Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision affect theory: Emotional reactions to the outcomes of risky options. *Psychological Science*, 8, 423–429.
- Mellers, B. A., Schwartz, A., & Ritov, I. (1999). Emotion-based choice. *Journal of Experimental Psychology: General*, 128, 332–345.
- Miller, D. T. (2001). Disrespect and the experience of injustice. *Annual Review of Psychology*, 52, 527–553.
- Miller, D. T., & Turnbull, W. (1990). The counterfactual fallacy: Confusing what might have been with what ought to have been. *Social Justice Research*, 4, 1–19.
- Morris, J. H., & Moberg, D. J. (1994). Work organizations as contexts for trust and betrayal. In T. Sarbin, R. Carney, & C. Eoyang (Eds.), *Citizen espionage: Studies in trust and betrayal* (pp. 163–187). Westport, CT: Praeger.
- Morrison, E. W., & Robinson, S. L. (1997). When employees feel betrayed: A model of how psychological contract violation develops. *Academy of Management Review*, 22, 226–256.
- National Highway Traffic Safety Administration. (1999, June 14). Most shorter drivers can eliminate risk without deactivation. [Online]. Available: <http://www.nhtsa.dot.gov/>.
- Nunnally, J. (1978). *Psychometric theory*. New York: McGraw Hill.
- Ordóñez, L. D., & Connolly, T. (2000). Comment: Regret and responsibility: A reply to Zeelenberg et al., 1998. *Organizational Behavior and Human Decision Processes*, 81, 132–142.
- Oxford English Dictionary, second ed. (1989). Prepared by J. A. Simpson and E. S. C. Weiner. Oxford: Clarendon Press.
- Parker-Pope, T. (2000). Headache sufferers learn aspirins cause more pain. *Wall Street Journal*, B1.
- Pear, R. (1998). Preventative use of Tamoxifen is allowed. *New York Times*, A22.
- People v. Lowery 894 P. 2d 758 (S.Ct. CO 1995).
- Polinsky, M., & Shavell, S. (1999). Punitive damages: An economic analysis. *Harvard Law Review*, 111, 869–962.
- Ritov, I. (1994). Can goals be uniquely defined? *Behavioral and Brain Sciences*, 17, 28–29.
- Ritov, I., & Baron, J. (1990). Reluctance to vaccinate: Omission bias and ambiguity. *Journal of Behavioral Decision Making*, 3, 263–277.
- Ritov, I., & Baron, J. (1999). Protected values and omission bias. *Organizational Behavior and Human Decision Processes*, 79, 79–94.
- Robinson, S. L. (1996). Trust and breach of the psychological contract. *Administrative Science Quarterly*, 41, 574–599.
- Robinson, S. L., & Morrison, E. W. (2000). The developments of psychological contract breach and violation: A longitudinal study. *Journal of Organizational Behavior*, 21, 525–546.
- Robinson, S. L., & Rousseau, D. M. (1994). Violating the psychological contract: Not the exception but the norm. *Journal of Organizational Behavior*, 15, 245–259.
- Rose, C. M. (1995). Trust in the mirror of betrayal. *Boston University Law Review*, 75, 531–558.
- Rothbart, M., & Park, B. (1986). On the confirmability and disconfirmability of trait concepts. *Journal of Personality and Social Psychology*, 50, 131–142.
- Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, 23, 393–404.
- Scallen, E. A. (1993). Promises broken vs. promises betrayed: Metaphor, analogy and the new fiduciary principle. *University of Illinois Law Review*, 897–980.
- Simonson, I. (1992). The influence of anticipating regret and responsibility on purchase decisions. *Journal of Consumer Research*, 19, 105–118.
- Slovic, P. (1987). Perception of risk. *Science*, 236, 280–285.
- Slovic, P. (1997). Public perception of risk. *Journal of Environmental Health*, 59, 22–54.
- Slovic, P. (2000). Rational actors and rational fools: The influence of affect on judgment and decision-making. *Roger Williams University Law Review*, 6, 163–212.
- Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology*, 28, 1–39.
- Slovic, P., Fischhoff, B., & Lichtenstein, S. (1982). Facts and fears: Understanding perceived risk. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 463–489). New York: Cambridge University Press.
- Slovic, P., Flynn, J., & Gregory, R. (1994). Stigma happens: Social problems in sting of nuclear waste facilities. *Risk Analysis*, 14, 773–777.
- Snepp v. US 444 US 507, 100 S. Ct. 763, 62 L. Ed.2d 704, 1980.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological methodology* (pp. 290–293). Washington: American Sociological Association.
- Strauss, M. B. (1994). *Violence in the lives of adolescents*. New York: W.W. Norton.
- Sunscreen takes the heat, 1998, February 24. Chicago Tribune, p. C3.
- Sunstein, C. R., Schkade, D., & Kahneman, D. (2000). Do people want optimal deterrence? *Journal of Legal Studies*, 29, 237–254.
- Switching off the air bag (1997, December 1). US News and World Report, p. 41.
- Teigen, K. H., Brun, W., & Frydenlund, R. (1999). Judgments of risk and probability: The role of frequentist information. *Journal of Behavioral Decision Making*, 12, 123–139.
- Tetlock, P. E. (1994). The consequences of taking consequentialism seriously. *Behavioral and Brain Sciences*, 17, 31–32.
- Tetlock, P. E., Peterson, R. S., & Lerner, J. S. (1996). Revising the value pluralism model: Incorporating social content and context postulates. In C. Seligman, J. M. Olson, & M. P. Zanna (Eds.), *The psychology of values: The Ontario Symposium* (Vol. 8, pp. 25–51).
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453–458.
- Tyler, T. R., & Boeckmann, R. (1997). Three strikes and you are out, but why? *Law and Society Review*, 31, 237–265.
- Vlek, C., & Stallen, P. J. (1981). Judging risks and benefits in the small and in the large. *Organizational Behavioral and Human Decision Processes*, 28, 235–271.
- Wald, M. L. (1996). The dangers from air bags and ways to avoid them. *New York Times*, C2.
- Wald, M. L. (1997). Preaching caution, officials allow cutoff switches for air bags. *New York Times*, A17.

- Wandersman, A. J., & Hallman, W. K. (1993). Are people acting irrationally? Understanding public concerns about environmental threats. *American Psychologist*, *48*, 681–686.
- Zeckhauser, R. J., & Viscusi, W. K. (1990). Risk within reason. *Science*, *248*, 559–564.
- Zeelenberg, M. (1999). Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making*, *12*, 93–106.
- Zeelenberg, M. J., Beattie, J., van der Pligt, J., & de Vries, N. K. (1996). Consequences of regret aversion: Effects of expected feedback on risky decision making. *Organizational Behavior and Human Decision Processes*, *65*, 148–158.
- Zeelenberg, M., van Dijk, W. W., van der Pligt, J., Manstead, A. S. R., van Empelen, P., & Reinderman, D. (1998). Emotional reactions to the outcomes of decisions: The role of counterfactual thought in the experience of regret and disappointment. *Organizational Behavior and Human Decision Processes*, *75*, 117–141.

Received 24 July 2001