Justifying One's Transgressions: How Rationalizations Based on Equity, Equality, and Need Affect Trust After Its Violation

Peter H. Kim and Derek J. Harmon
University of Southern California

We investigate how efforts to justify a transgression as an attempt to address matters of equity, equality, or need would affect the implications of an apology for trust after its violation, and how this would depend on the intended beneficiary. To do so, we conducted 2 studies, including a new research design that supplements the rigor of experiments with far greater realism. Although combining a justification with an apology tended to elicit higher trust relative to an apology alone when the violation benefited another party, doing so was ineffective or harmful when the violation benefited the violator. Finer-grained analyses comparing the 3 types of justifications, furthermore, revealed that the addition of equity-based justifications elicited higher trust than the addition of equality- or need-based justifications in general, and that the addition of need-based justifications was particularly harmful when the violation benefited the self. Perceived fairness mediated these effects.

Keywords: trust, violation, apology, justification, beneficiary

Supplemental materials: http://dx.doi.org/10.1037/xap0000030.supp

Although trust has been widely recognized to offer important benefits for both social and organizational life, numerous examples from the research literature and the popular press have revealed that trust is also frequently violated (e.g., Conway & Briner, 2002; Robinson & Rousseau, 1994). For this reason, scholars have begun to consider the implications of a wide range of verbal and substantive responses that might be used to address such incidents (e.g., Bottom, Gibson, Daniels, & Murnighan, 2002; Kramer & Lewicki, 2010; Lount, Zhong, Sivanathan, & Murnighan, 2008; Schweitzer, Hershey, & Bradlow, 2006). These studies have provided many insights regarding when and why these different responses might be effective. However, they have largely focused on dyadic settings rather than settings in which additional parties might be involved.

This oversight represents a major shortcoming in the literature for at least two reasons. First, the neglect of multiparty contexts ignores the fact that individuals are often highly interconnected, for example, because of their mutual reliance on limited resources, rules and procedures that involve the input of multiple constituencies, as well as their broader networks of relationships. Second, even when those other parties are not actively involved in a given situation, this interconnectedness may increase both the potential motivations for the violation and the ways in which the violator might subsequently address the incident. More specifically, it raises the possibility that a

violator might have sought to benefit either the self or another, and have done so not only for the sake of the intended beneficiary but also to address a range of other goals. These notions, in turn, highlight the need to investigate how one might address such incidents, not simply by accepting more or less responsibility for the violation, but by altering perceptions of the violation itself, specifically through different types of justifications.

Imagine, for example, you discovered that someone had tampered with the admissions process at a university in order to admit a new applicant. How would you react if that person claimed to have done so to address a source of unfairness in the admission system? More specifically, would your willingness to trust this person in the future be greater if that party simply apologized for the incident or combined that apology with a justification that (a) the applicant's accomplishments had not been adequately recognized, (b) the applicant had not been given the same opportunity as others to succeed, or (c) the applicant really needed to attend that university? Moreover, how might your reactions differ if the tampering had concerned the violator's own application as opposed to another applicant?

Although the potential value of justifications has been suggested by the broader literature on social accounts (Scott & Lyman, 1968) and distributive justice (Deutsch, 1975), as well as recent trust repair theory (Kim, Dirks, & Cooper, 2009), their implications for trust after its violation remains empirically unexplored. Thus, we still know quite little about how justifications would affect trust after such incidents and under what conditions this is more likely to occur. The purpose of this inquiry is to begin addressing these limitations by investigating how three of the most fundamental types of justifications that have been underscored by the distributive justice literature (equity, equality, and need) would affect the implications of an apology for trust, specifically after an integrity-based violation, and how those relationships may depend on whether the intended beneficiary had been the self or another.

Trust and Trust Violations

Trust is defined as a psychological state comprising the intention to accept vulnerability based on positive expectations of the intentions or behavior of another (Rousseau, Sitkin, Burt, & Camerer, 1998). For clarity, throughout this article, we use the term "trustees" to refer to those whose trustworthiness is being evaluated, and the term "trustors" to refer to those evaluating the trustee. Further, with regard to what have been considered two of the most important determinants of trust, researchers have defined "competence" as the degree to which the trustee possesses the technical and interpersonal skills required for a job, and "integrity" as the degree to which the trustee adheres to a set of principles that the trustor finds acceptable (e.g., Mayer, Davis, & Schoorman, 1995). Hence, consistent with past research (e.g., Kim, Ferrin, Cooper, & Dirks, 2004), we distinguish competence- from integrity-based violations by whether the transgression had been unintentional (based on a lack of requisite skills) or intentional (based on a disregard of principles the trustor finds acceptable), respectively.

Finally, it is critical to note that although people can often behave in ways that violate trust, such as by exploiting trustors' dependencies or by neglecting to fulfill their expectations, research suggests that trust can be violated even by unsubstantiated allegations (Bell & Loftus, 1989). Moreover, evidence from these and other studies have shown that trust can be violated even with those who have not been directly harmed by the transgression (e.g., Ferrin, Kim, Cooper, & Dirks, 2007; Kim et al., 2004), and that such observers are often willing to punish violators as well (Fehr & Fischbacher, 2004; Nelissen & Zeelenberg, 2009). For these reasons, the need to address such incidents can involve not just victims but also bystanders, as well as those only hearing about such allegations well after the fact.

Justifying Integrity-Based Violations

Research on addressing trust violations has revealed that the relative benefits of such responses can be highly contingent on the type of violation. Whereas apologies (Kim, Cooper, Dirks, & Ferrin, 2013; Kim et al., 2004), efforts to assume more blame with an internal attribution (Kim, Dirks, Cooper, & Ferrin, 2006; Tomlinson, Dineen, & Lewicki, 2004), and substantive acts such as penance, regulation, and reparation (Desmet, De Cremer, & Van Dijk, 2011; Dirks, Kim, Ferrin, & Cooper, 2011), for example, have been found to be relatively beneficial for trust when violations concern matters of competence, they are either ineffective or harmful for trust when violations concern matters of integrity. Indeed, for integrity-based violations, studies have found higher levels of trust after trustees instead denied culpability for the transgression (Ferrin et al., 2007; Kim et al., 2004, 2013). The obvious problem with simply denying culpability for integrity-based violations, however, is that trustees might actually have committed them. In such cases, it would not only be unethical to deny culpability but also unconvincing when there is evidence of guilt (Kim et al., 2004).

One solution the literature has offered for addressing this problem is for trustees to acknowledge that they committed the violation, but then provide an explanation for the untoward behavior. Explanations such as excuses, which concern efforts to explain that the act was at least partly caused by external forces, and justifications, which concern efforts to claim that the act was actually appropriate by pointing to the fulfillment of some superordinate goal, have been found to affect perceptions of justice, as well as behaviors such as cooperation, retaliation, and withdrawal (see Shaw, Wild, & Colquitt, 2003, for a review). Moreover, recent theory on trust repair suggests that excuses and justifications may elicit higher levels of trust relative to simple acknowledgments of guilt (Kim et al., 2009). Yet although evidence supports the notion that excuses can elicit higher trust than efforts to assume full blame under certain conditions (Kim et al., 2006), the implications of justifications in general for trust after its violation, let alone justifications of different types, have remained empirically unexplored.

The importance of comparing the implications of different types of justifications has been well established by research on distributive justice. Distributive justice concerns the distribution of the conditions and goods that affect individual well-being (Deutsch, 1975), and much of the research in this domain has focused on how perceptions of distributive justice might be achieved by adhering to three alternative allocation principles—equity, equality, and need. Indeed, Deutsch (1975) observed that these specific principles represent three of the most basic, common, and substantive values in resource allocation situations. Whereas equity concerns whether resources are allocated in proportion to contributions or inputs, equality concerns whether people receive the same benefits or have the same opportunity to benefit, and need concerns whether people are rewarded based on their level of need or deprivation (Conlon, Porter, & Parks, 2004).

These three principles have been found to play an important role in determining perceptions of distributive justice (e.g., Austin, 1980; Conlon et al., 2004; Mannix, Neal, & Northcraft, 1995; Reis & Gruzen, 1976), which has, in turn, been found to be positively correlated with trust (Cohen-Charash & Spector, 2001; Colquitt, Conlon, Wesson, Porter, & Ng, 2001). Moreover, because evidence suggests that violating these distributive justice principles can lower trust (e.g., Stouten, De Cremer, & Van Dijk, 2006), one might also expect that justifications that concern efforts to maintain these principles would prove beneficial for trust after a violation has occurred. More specifically, to the extent that a trustee is believed to have affected people's outcomes in an unwarranted manner, trustors are likely to consider the incident unfair (Adams, 1965). This sense of unfairness, in turn, has been observed to lower trustors' sentiments about the trustee's trustworthiness and reduce their willingness to engage with that party (Lazarus & Cohen-Charash, 2001; i.e., their trust in that party would be lowered). Yet the trustee may subsequently be able to alter the extent to which the trustor considers the behavior unfair by justifying the behavior in some way. In particular, the trustee might explain that their behavior was indeed fair because it attempted to promote a distributive justice goal, such as equity, equality, or need. To the extent that trustors value the goal underscored by the trustee's justification, trustors should consider the incident fairer than they might have initially presumed, and thereby exhibit higher trust toward that party than if the justification had not been conveyed.

Potential Complications

However, at least two problems may complicate an extension of findings regarding these three distributive rationales from the distributive justice literature to the trust violation context. First,

whereas distributive justice tends to involve systemic considerations, specifically, of how allocations should occur across all members of a group or community, trust violations tend to be more targeted. Particularly in cases of integrity-based violations, they typically concern efforts to benefit or harm just one or a few individuals, irrespective of (or with far less consideration of) the implications of that violation for others. In such cases, the choice of target (i.e., the specific individual[s] to be directly benefited or harmed by the violation) may have some effect on the implications of these distributive justice rationales. In particular, evidence suggests that people's judgments of fairness are typically biased in favor of their self-interests (Babcock, Loewenstein, Issacharoff, & Camerer, 1995; Messick & Sentis, 1979). Hence, trustors may recognize this tendency and thereby respond less favorably to their trustee's justifications if the violation had been intended to benefit the trustee personally rather than another party or the group more generally. If so, one might expect that the same justification ultimately would be less beneficial for trust when the violation had been intended to benefit the trustee rather than another party, given the greater concern for collective interests the latter would entail.

Second, it is unclear whether the relative benefits of equity, equality, and need rationales for perceptions of distributive justice would also be found after trust violations. The distributive justice literature has observed that people seek commensurability ("fit") between the primary goals of a relationship and the resource distribution rules used in that relationship (Mannix et al., 1995). Moreover, researchers have consistently observed that equity rather than equality or need tends to be the dominant principle of distributive justice in relationships in which economic productivity is the primary goal (Conlon et al., 2004; Deutsch, 1975; Mannix et al., 1995). This observation is based on the notion that people in task-directed relationships tend to favor rules that allocate scarce resources between different members in ways that maximize overall productivity and efficiency within the relationship. These rules are oriented toward discovering differences among members' contributions to encourage and reward the contributions of the most able, and they are driven by the goal of task achievement (Deutsch, 1985, pp. 88–91), given that it could otherwise be difficult to fulfill the primary purpose for which these members had assembled (i.e., economic productivity). Thus, to the extent that our research context is oriented toward performance, and this does not change after a trust violation has occurred, the relative importance placed on equity over equality or need for addressing distributive justice concerns should ultimately lead equity-based justifications to elicit higher levels of trust after its violation than justifications based on equality or need.

Yet this relative efficacy of equity-, equality-, and need-based justifications may also depend on the intended beneficiary. In particular, if the implications of these justifications for trust are based on their ability to recast the violation as an attempt to fulfill broader collective interests (Kim et al., 2009), then their effects should hinge not only on what those collective interests might be (e.g., the aforementioned tendency to emphasize equity over equality or need in productivity-oriented relationships) but also on whether the justifications lead trustors to believe that the collective had actually been considered. Otherwise, it would be difficult to support the claim that distributive justice across the collective had been a concern. This perception that the collective had been considered can readily be conveyed by all three rationales when

the intended beneficiary is another party, because this makes the violator's concern with other members of the collective explicit. Further, even when the intended beneficiary is the self, both equity- and equality-based rationales inherently involve at least some consideration of others, given that this is required to determine whether equity or equality across the collective has been achieved, respectively (Conlon et al., 2004). However, the implications of need-based rationales may be quite different depending on whether the intended beneficiary is another party or oneself. Whereas a need-based rationale clearly suggests that the collective had been considered when the intended beneficiary is another party, this response may seem particularly self-serving when the intended beneficiary is the self and, thereby, fail to suggest that the collective had been considered at all. If so, need-based justifications may be especially ineffective for (a) recasting the violation as an attempt to fulfill broader collective interests; (b) persuading trustors that the trustee had, thus, been fairer than they had initially presumed; and, ultimately, (c) eliciting higher levels of trust, relative to equity- and equality-based justifications, when the intended beneficiary is the self rather than another party.

Predictions

These considerations lay the foundation for several hypotheses regarding the manner in which justifications would affect the implications of apologies for trust after its violation. If apologizing in a manner that claims that the act was actually appropriate, by pointing to the fulfillment of some superordinate goal, can elicit higher trust than apologies that accept full blame after an integrity-based trust violation (Kim et al., 2009), one might expect that combining an apology with a justification would generally produce higher levels of trust than an apology alone after an integrity-based violation has occurred. However, the aforementioned considerations also suggest that justifications may prove less effective when the violation had been intended to benefit the self rather than another party. Thus, we can ultimately integrate these notions to predict an interaction between the intended beneficiary (self vs. other) and type of response (apology plus justification vs. apology alone), in which an apology that has been combined with a justification would elicit higher trust (than an apology alone) when the violation was intended to benefit another party (than when the violation was intended to benefit the self).

> *Hypothesis 1:* An apology combined with a justification will result in higher trust (than an apology alone) when the violation was intended to benefit another party (than when the violation was intended to benefit the self).

Finer-grained considerations, furthermore, suggest two additional hypotheses concerning the relative benefits of combining an apology with these different types of justifications. To the extent that people's general tendencies to prioritize equity over equality or need in performance-oriented relationships can be generalized to the trust-violation context, equity-based justifications should be more effective than equality- or need-based justifications at persuading trustors that the violation might have actually been fair. If so, this should ultimately allow the combination of an apology with an equity-based justification to elicit higher levels of trust than the combination of an apology with a justification based on equality or need.

Hypothesis 2: An apology combined with an equity-based justification will result in higher trust than an apology combined with a justification based on equality or need.

However, to the extent that the relative benefits of these rationales for trust further hinge on whether they can recast the violation as an attempt to fulfill broader collective interests (specifically, to improve distributive justice), one might expect that this will also depend on the extent to which they lead trustors to believe that the collective had been considered in some way. This is particularly unlikely to occur when need-based rationales are offered for violations intended to benefit the self (i.e., because they neither benefit another party nor rely on principles, such as equity or equality, that inherently require comparisons with others). For this reason, the combination of an apology with a need-based justification when the intended beneficiary is the self should result in lower levels of trust than the combination of an apology with an equity- or equality-based justification when the intended beneficiary is the self or the combination of an apology with any of these justifications (equity-, equality-, or need-) when the intended beneficiary is another.

Hypothesis 3: An apology combined with a need-based justification when the intended beneficiary is the self will result in lower trust than an apology combined with an equity- or equality-based justification when the intended beneficiary is the self or an apology combined with equity-, equality-, or need-based justifications when the intended beneficiary is another.

Finally, given that (a) these equity, equality, and need rationales would seek to address trust violations by recasting the transgression as an attempt to improve distributive justice; (b) distributive justice has been observed to affect perceptions of fairness (Adams, 1965); and (c) perceived fairness has been found to be an important determinant of trust (see meta-analyses by Cohen-Charash & Spector, 2001, and Colquitt et al., 2001), one might expect that perceived fairness will mediate these effects. If so, we should observe a causal sequence in which these distributive justice rationales affect the perceived fairness of the incident, and this perception of fairness ultimately affects trust.

Hypothesis 4: Perceived fairness will mediate the interactive effect of intended beneficiary (self vs. other) and violation response (apology plus justification vs. apology alone) on trust.

Study 1

To begin investigating these hypotheses, we first conducted an online experiment with a 2 (intended beneficiary: self vs. other) × 4 (violation response: apology plus equity vs. apology plus equality vs. apology plus need vs. apology alone) between-subjects factorial design.

Method

Participants. Four hundred forty-seven participants were recruited from Amazon's Mechanical Turk (MTurk) and were each paid $2 for their involvement. Participants were randomly assigned to the eight study conditions. The participants averaged 32.3

($SD = 10.9$) years of age, 8.9 years of work experience ($SD = 9.1$), and 57% were male. With regard to ethnicity, 65% of these participants were Caucasian, 19% were South Asian, 6% East Asian, 6% were Hispanic, and 4% were African American.

Procedure. The study asked participants to imagine that they worked for a 20-person sales division of a company for which they were involved in a sales competition last year. The way the competition worked was that the salespersons with the five highest sales totals for a specified month would receive company-wide recognition for their performance. Because calculating all the sales numbers was a long process, after the competition was over, the company asked for one of the salespeople to volunteer to add up and report all the sales numbers to the manager. However, before the winners from the competition were announced, the manager who received the final sales numbers from the volunteer alleged that this salesperson changed one of the final sales numbers to be different from what it should have been. This prompted an investigation by the company's human resource department. After reading this vignette, participants were given the manager's allegation and the accused salesperson's response and were then asked to answer several questions.

Manipulations.

Intended beneficiary. The trust violation was described as having been committed to benefit either the volunteer salesperson or another salesperson. In both conditions, the volunteer salesperson was accused of intentionally changing the intended beneficiary's total sales, while correctly tallying all of the other scores.

Violation response. The accusation was then followed by a response from the volunteer salesperson, in which this party either simply apologized for the act in question or combined that apology with a justification. The combined response was used, rather than a justification in which an apology was not conveyed, because both apologies and justifications have been found to be comprised of multiple components that typically overlap one another. An expression of one's responsibility for the offense, for example, has been found to be an inherent component of both apologies (Scher & Darley, 1997) and justifications (Kim et al., 2009), and justifications are commonly conveyed in a manner that communicates other components of an apology as well, such as an expression of regret for the harm the offense might have caused others and a promise of forbearance. This may be done to emphasize that the offense had been an exception that only arose from the particulars of the situation and as a social lubricant to acknowledge the inconvenience or harm the violation might have caused others (Maddux, Kim, Okumura, & Brett, 2011) in order to improve the favorability of trustors' subsequent evaluations. Indeed, one might expect that this would be particularly important in situations such as those in the present research, in which the trustee might face negative repercussions. Thus, for these reasons, it seemed more reasonable and more realistic to combine the trustee's justification with an apology, rather than have the response suggest the trustee did not regret the offense at all.

This comparison of simple apologies with apologies combined with justifications entails that the apology itself should ultimately be considered a constant feature of this study context, against which the implications of adding the three kinds of justifications can be compared. In the apology condition, participants read, "I am truly sorry, what the manager said about me intentionally changing [my/another salesperson's] sales to be higher than what [I/they]

actually got is true. Again I'm really sorry." Each of the apology plus justification conditions then provided an additional sentence that read, "I don't want to get into the details, but I did this to address a matter of [inequity/inequality/need]."

Measures.

Trusting intentions. Five items assessing participants' trusting intentions were adapted from Kim et al. (2004)—(a) "I would feel comfortable working with this individual in the future," (b) "I would feel comfortable with this individual working on a similar task in the future without oversight," (c) "If I worked with this individual again, I would keep my eye on them (reverse-scored)," (d) "If I had a choice, I wouldn't let this individual have any influence over issues that are important to me" (reverse-scored), and (e) "I would give this individual a task that was critical to me, even if I could not monitor their actions"—using a 7-point Likert scale (1 = *strongly disagree*; 7 = *strongly agree*; $M = 2.50$, $SD = 1.20$, $\alpha = .84$).

Perceived fairness. Two items assessing participants' perceptions of the volunteer's fairness were drawn from Conlon et al. (2004)—(a) "In my opinion, the way this individual scored these tests seems fair," and (b) "This individual's reasoning was fair"—using a 7-point Likert scale (1 = *strongly disagree*; 7 = *strongly agree*; $M = 2.62$, $SD = 1.58$, $\rho = .87$).[1]

Pretest

Prior to conducting the main study, we first sought to assess the suitability of this context for studying trust violations and the effectiveness of our manipulations with a two-phase pretest of 43 MTurk participants who were each paid $1 for their involvement. The participants averaged 30.16 ($SD = 7.77$) years of age, 7.0 years of work experience ($SD = 7.2$), and 73% were male. Moreover, 66% of these participants were Caucasian, 25% were South Asian, 6% were East Asian, and 3% chose "Other." None of the participants failed any of the manipulation- or attention-check questions.

The first phase of this pretest sought to verify that trust would be violated as a result of this transgression with a 2 (intended beneficiary: self vs. other) × 2 (time: previolation vs. postviolation) repeated measures design that assessed participants' trusting intentions toward the trustee twice after reading about the study context. Specifically, we collected this measure immediately before participants learned about the transgression and then immediately after learning about the transgression but prior to the trustee's response. This analysis revealed that trusting intentions toward the trustee significantly dropped from before knowing about the transgression ($M_{self} = 4.23$, 95% confidence interval [CI] [3.84, 4.62]; $M_{other} = 4.18$, 95% CI [3.79, 4.57]) to after ($M_{self} = 2.37$, 95% CI [1.98, 2.76]; $M_{other} = 2.37$, 95% CI [1.90, 2.84]), $F(1, 42) = 77.48$, $p < .001$, $\eta^2 = .65$, 95% CI [1.43, 2.25]. Moreover, this drop in trust did not significantly differ based on whether the intended beneficiary was the self or another, as evidenced by the interaction failing to reach significance, $F(1, 42) = .02$, $p = .89$, $\eta^2 = .00$.

Next, to verify that participants understood what the terms equity, equality, and need entailed, the second phase of this pretest presented each participant with one of the three justifications and asked them to select the definition (based on Conlon et al., 2004) that best captured this fairness principle. Our findings revealed that 14 out of 15 participants correctly defined equity, 15 out of 15 participants correctly defined equality, and 14 out of 14 participants correctly defined need, providing strong support for the notion that participants understood and could distinguish among these terms.

Results

Manipulation checks for the main study revealed that 98% of participants correctly identified the intended beneficiary and 96% identified the correct response. We excluded the participants that missed one or more of these questions from the sample. Confirmatory factor analyses, furthermore, indicated an adequate fit and supported convergent validity for a two-factor model in this study that included Trusting Intentions and Perceived Fairness ($r = .76$), $\chi^2(12, N = 447) = 41.05$, Comparative Fit Index (CFI) = .98, Normed Fit Index (NFI) = .98, Tucker-Lewis Index (TLI) = .97, Root Mean Square Error of Approximation (RMSEA) = .074, all item-factor loadings ≥.47 ($p < .001$). Finally, discriminant analyses (Bagozzi & Phillips, 1982) indicated that the hypothesized two-factor model fit the data significantly better than a one-factor model, $\chi^2(13, N = 447) = 81.09$, $\Delta\chi^2(1, N = 447) = 40.04$, $p < .001$. Table 1 reports variable means and standard deviations by condition.

To assess how justifications would affect the implications of apologies for Trusting Intentions and how this might depend on the intended beneficiary, we first conducted a 2 (Intended Beneficiary: self vs. other) × 4 (Violation Response: apology plus equity vs. apology plus equality vs. apology plus need vs. apology alone) ANOVA. This analysis revealed a significant main effect for Intended Beneficiary, $F(1, 439) = 16.22$, $p < .001$, $\eta^2 = .04$, 95% CI [.23, .67], a significant main effect for Violation Response, $F(3, 439) = 5.62$, $p < .01$, $\eta^2 = .04$, and a significant two-way interaction, $F(3, 439) = 2.86$, $p < .05$, $\eta^2 = .02$.

Then, to Test Hypothesis 1, which predicted that an apology that has been combined with a justification would result in higher trust (than an apology alone) when the violation was intended to benefit another party (than when the violation was intended to benefit the self), we conducted a planned contrast of this interaction effect (Contrast #1). This involved coding each of the apology plus justification conditions as 1, coding the apology alone condition as −3, and then testing how this violation response comparison interacted with Intended Beneficiary (with "self" coded as 1 and "other" coded as −1) to affect Trusting Intentions. This analysis revealed a significant two-way interaction, $F(1, 439) = 6.80$, $p < .01$, $\eta^2 = .02$, whereby an apology plus justification resulted in higher Trusting Intentions than an apology alone when the violation was intended to benefit another salesperson, $F(1, 219) = 13.05$, $p < .001$, $\eta^2 = .06$, but this is not the case when the violation was intended to benefit the self, $F(1, 220) = 0.02$, $p = .98$, $\eta^2 = .00$, supporting Hypothesis 1 (see Figure 1).

We then conducted finer-grained comparisons to evaluate how the combination of an apology with justifications based on equity, equality, and need would compare with one another with two additional planned contrasts that we had developed to evaluate Hypotheses 2 and 3. To evaluate Hypothesis 2, which predicted

---

[1] The Spearman-Brown statistic is used to calculate reliability when a scale only has two items (Eisinga, Grotenhuis, & Pelzer, 2013).

Table 1
Study 1: Means, Standard Deviations, and CIs by Condition

	M	*SD*	95% CI	*M*	*SD*	95% CI
Justification (plus apology)	Self (*n* = 174)			Other (*n* = 169)		
Trusting intentions	2.28	1.10	[2.12, 2.44]	3.89	1.25	[3.70, 4.08]
Perceived fairness	2.24	1.31	[2.05, 2.43]	3.26	1.70	[3.00, 3.52]
Apology (alone)	Self (*n* = 50)			Other (*n* = 54)		
Trusting intentions	2.28	1.21	[1.94, 2.62]	2.21	1.00	[1.94, 2.48]
Perceived fairness	2.28	1.67	[1.82, 2.74]	2.09	1.24	[1.76, 2.42]
Equity	Self (*n* = 57)			Other (*n* = 55)		
Trusting intentions	2.52	1.11	[2.23, 2.81]	3.05	1.39	[2.68, 3.42]
Perceived fairness	2.48	1.42	[2.11, 2.85]	3.42	1.77	[2.95, 3.89]
Equality	Self (*n* = 57)			Other (*n* = 59)		
Trusting intentions	2.43	1.25	[2.11, 2.75]	2.88	1.07	[2.61, 3.15]
Perceived fairness	2.44	1.31	[2.10, 2.78]	3.36	1.60	[2.95, 3.77]
Need	Self (*n* = 60)			Other (*n* = 55)		
Trusting intentions	1.90	0.82	[1.69, 2.11]	2.74	1.29	[2.40, 3.08]
Perceived fairness	1.83	1.10	[1.55, 2.11]	3.00	1.73	[2.54, 3.46]

Note. CI = Confidence Interval.

that the addition of an equity-based justification to an apology would result in higher trust than the addition of an equality- or need-based justification, in general, Contrast #2 coded the equity conditions as -2, the equality and need conditions as 1, and the apology alone conditions as 0. Analysis of this contrast revealed that the addition of an equity-based justification to an apology resulted in significantly higher Trusting Intentions than the addition of equality- or need-based justifications, $F(1, 439) = 4.99$, $p < .05$, $\eta^2 = .01$, supporting Hypothesis 2.

Next, to evaluate Hypothesis 3, which predicted that an apology combined with a need-based justification when the intended beneficiary is the self would result in lower trust than an apology combined with an equity- or equality-based justification when the intended beneficiary is the self or an apology combined with any

Figure 1. Study 1: Effects of an apology versus justification as a function of the intended beneficiary.

of these justifications (equity, equality, or need) when the intended beneficiary is another, we developed Contrast #3 based on the recommendations of Buckless and Ravenscroft (1990). Specifically, given that one apology plus justification condition (i.e., the need-based justification when the intended beneficiary is the self) was predicted to differ from the others, the first condition was coded as 5, whereas the other apology plus justification conditions were coded as -1, and the apology alone conditions were coded as 0. By doing so, this contrast enabled a more precise test of Hypothesis 3's predicted interaction than the interaction term standard statistical packages would use by default. This analysis revealed that the combination of an apology with a need-based justification when the intended beneficiary is the self led to significantly lower Trusting Intentions than the combination of an apology with an equity- or equality-based justification when the intended beneficiary was the self or the combination of an apology with any of the justifications when the intended beneficiary was another party, $F(1, 444) = 24.97$, $p < .001$, $\eta^2 = .04$, supporting Hypothesis 3 (see Figure 2).

Mediation Analyses

We then tested Hypothesis 4's prediction that perceived fairness would mediate the Intended Beneficiary × Violation Response interaction's effects on Trusting Intentions by using the aforementioned comparison between the apology plus justification conditions (each coded as 1) versus the apology alone condition (coded as -3) as the independent variable, the intended beneficiary (with "self" coded as 1 and "other" coded as -1) as the moderator, and bootstrapping (with 5,000 samples) the indirect effects of this interaction on Trusting Intentions through Perceived Fairness (Preacher & Hayes, 2008; see Figure 3). More specifically, we tested the moderated mediation prediction using Hayes' PROCESS macro (Model 7), and used the index of moderated
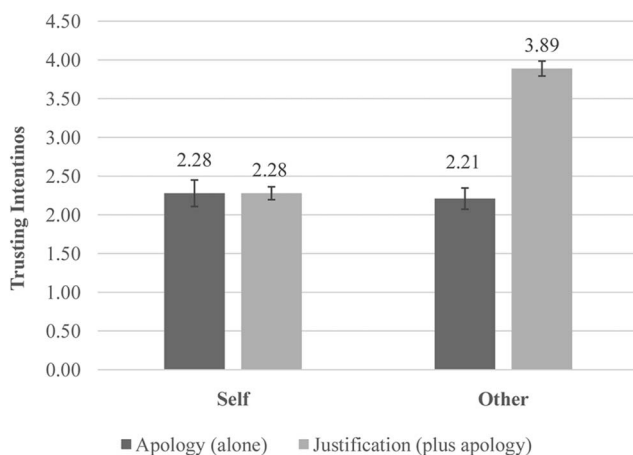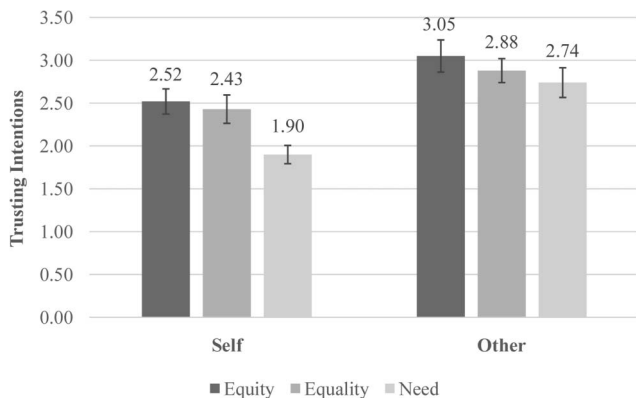
Figure 2. Study 1: Effects of equity, equality, and need as a function of the intended beneficiary.

mediation to evaluate the model's significance. The results revealed a significant indirect effect as evidenced by the fact that the 95% CI did not include zero [−.27, −.09], thus supporting Hypothesis 4 (see Table 2).

Discussion

Study 1 demonstrates strong support for our predictions. An apology plus justification resulted in higher Trusting Intentions than an apology alone when the violation was intended to benefit another party but not when the violation was intended to benefit the self. Moreover, this relationship was mediated by perceived fairness. Finer-grained comparisons also revealed that the addition of equity-based justifications to an apology resulted in higher Trusting Intentions than the addition of equity- and need-based justifications, in general, and that the addition of need-based justifications was particularly detrimental when the violation had been intended to benefit the self. Thus, given this initial support for the hypotheses, we developed a second study to bolster and extend these conclusions.

Study 2

The purpose of this study was to extend our previous findings in several ways. First, to ensure that the results were not limited to Study 1's specific context, we sought to replicate our findings in an entirely new setting. Second, we sought to verify that our findings would hold even if the violation concerned a different kind of outcome (i.e., a monetary award rather than minor recognition). Finally, we sought to enhance the psychological realism of this

research by investigating these hypotheses with participants who felt personally involved in a transgression they believed actually occurred, while maintaining the potential for experimental control and casual inferences.

To address these goals, we developed an experiment with a 2 (intended beneficiary: self vs. other) × 4 (violation response: apology plus equity vs. apology plus equality vs. apology plus need vs. apology alone) between-subjects factorial design and engaged in extensive additional efforts to bolster the perceived authenticity of the situation. In particular, we designed the study (a) to occur well beyond the physical and temporal confines of the ostensible experiment, (b) to be quite comparable with the kinds of honor code violations that have previously arisen and continue to be a real risk in this and similar participant populations, and (c) to be considered highly authentic in its potential implications for participants' actual outcomes. This study, as well as the others in this article, received full approval from our institution's human subjects review board.

Method

Participants. One hundred sixty-five undergraduate students enrolled in introductory organizational behavior courses participated in this study for course credit and were randomly assigned to the eight study conditions. They averaged 21 years of age ($SD = 3.1$), 2.6 years of work experience ($SD = 2.4$), and 59% were male. In terms of ethnicity, 42% of these participants were Caucasian, 38% were East Asian, 8% were Hispanic, 7% were South Asian, and 5% were African American.

Procedure. The structure of this study is similar to that of Study 1. Participants were placed in a competition in which another member commits a transgression, that member attempts to address the incident, and participants are then asked to evaluate that individual. Participants signed up online for a two-phase task competition study, wherein Phase 1 was an in-person session with up to a possible 24 other participants, and Phase 2 was an online survey to be completed at home a few days thereafter. Of the 10 sessions held, the average number of participants was 17 (minimum = 14; maximum = 23), with Phase 1 beginning on a Monday or Tuesday of the week and Phase 2 being completed no more than 5 days later. The variation in the number of participants per session arose from their occasional last-minute decisions not to participate because of personal conflicts, as well as the need to exclude participants that failed to finish the Phase 2 online survey within the allotted time frame. Phase 1 was identical for all sessions, whereas Phase 2 contained all the study manipulations. To minimize the risk of differences across the 10 sessions, we randomly
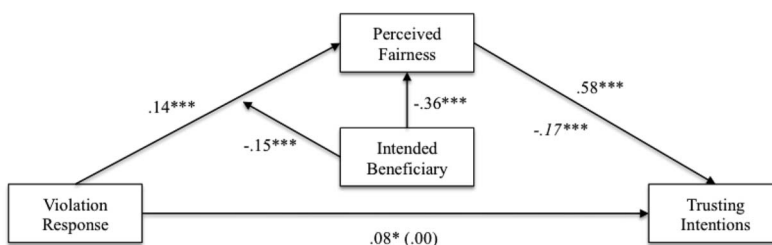

Figure 3. Study 1: Test of moderated mediation.

Table 2
Study 1: Results of Moderated Mediation

	B	SE	t	p	95% CI
Direct effects on fairness					
Violation response	0.14	0.04	3.37	0.000	[.06, .22]
Intended beneficiary	−0.36	0.07	−5.04	0.000	[−.50, −.22]
Violation Response × Intended Beneficiary	−0.15	0.04	−3.59	0.000	[−.23, −.07]
Direct effects on trusting intentions					
Violation response	0.08	0.03	2.50	0.013	[.02, .15]
Perceived fairness	0.58	0.02	24.56	0.000	[.53, .62]
Indirect effect on trusting intentions					
Violation Response × Intended Beneficiary → Perceived Fairness	−0.17	0.05	a	a	[−.27, −.09]

a. Statistic not applicable when assessing indirect effects and reporting the index of moderated mediation (Preacher & Hayes, 2008).

assigned each participant to one of the eight conditions in each session.

Phase 1 began with all the participants in the session meeting in a classroom with the researcher. After filling out consent forms, we told participants that the purpose of the study was to understand how people perform in a competition involving basic academic tasks while in the presence of others. We then informed participants that they would have 15 min to take a test containing 15 brain-teaser-type questions and that the top five scores from their session would be placed in an overall lottery for a chance to win one of three monetary prizes (gift cards of \$75, \$50, and \$25). Participants were also told that their tests were randomly generated from a question bank that contained questions of three difficulty levels (easy, medium, and hard), so they should not be surprised if some questions appeared more difficult than others. Though participants were told that they each had a unique test, in reality, they all had the exact same 15 questions, ordered in five different ways to give participants the impression they had a different test if they happened to glance at another student's materials. In actuality, the test had nothing to do with the study itself, although its format served as a basis for our later manipulations. Performance on the test was, however, used to determine the prize winners, as described during the session.

After the participants completed the test, the researcher collected the materials. The researcher then communicated to the participants that students have often expressed interest in learning about the research process, and that we would like to take this as an opportunity to ask for one volunteer from this session. Aside from this volunteer learning about this particular research project, the volunteer would also be asked to score the tests from this session the following morning and report the five top scores to their fellow session members (they would receive an e-mail list subsequent to grading the tests). In return, this volunteer would be exempt from Phase 2 and would receive full participation credit. The researcher asked participants to contact the researcher via e-mail immediately following Phase 1 if they were interested, and mentioned that if no one volunteered, one of them would be randomly selected. Participants were then thanked and notified that they should expect an e-mail from the researcher containing a survey for Phase 2 that would allow them to complete the study, as well as an e-mail from the volunteer reporting the top five scores from their session. Finally, the participants were dismissed.

In actuality, no volunteer was selected for any of the sessions. If a participant expressed interest in volunteering, they were provided a scripted response from the researcher, stating that more than one person had volunteered, that the final volunteer had been randomly selected from those interested, and that they had not been chosen.[2] As such, the participants always believed that one of the session members was scoring the tests, regardless of their own interest in doing so.

We also took several steps in preparation for Phase 2 to establish our cover story and maximize its perceived authenticity. First, we established a school e-mail address for the fictitious student volunteer, so this person would appear as a fellow student if participants searched for this student in the e-mail directory. Second, we established a fictitious ethics board called the Standards Violation in Research (SVR) Committee, which was ostensibly in charge of monitoring and dealing with issues related to the research subject pool. When research participants appeared to fail to meet the requested standards of research participation, the SVR Committee determined whether study credit should be granted and/or if a penalty should be imposed. We also developed a dedicated SVR e-mail account, SVR letterhead for communication, and a named contact individual on the committee. All communication, however, (whether from the fictitious volunteer or the fictitious SVR Committee) came from the researchers, without the participants' knowledge.

Phase 2 began with the participants expecting to receive an e-mail from the researcher with the final survey. Instead, they received a series of e-mail communications over the course of the following 4 days that served as our manipulations for all eight conditions. The first e-mail was sent from the researcher to all session participants the day after Phase 1 ended, shortly after the volunteer had apparently turned in the graded tests earlier that morning. This e-mail was intentionally vague, stating that when participating in studies, researchers expect that participants will conduct themselves with full effort, attention, and honesty. This e-mail alerted participants to the fact that something had occurred,

---

[2] Only three participants out of the 165 expressed a direct interest in volunteering, as a result of our deliberately offering only one opportunity to volunteer extremely early the next morning; controlling for these participants did not impact our results.

but did not provide specific information. Later that day, all session participants received a second e-mail from the fictitious volunteer, in which the individual acknowledged being the volunteer in their session and informed everyone in the session that (a) something came up during the scoring process, (b) that the volunteer was being penalized for this incident, and (c) that the volunteer was appealing this penalty to a committee that will likely be reaching out to those involved (i.e., everyone in the session) about this issue. The purpose of these e-mails was to gain the participants' attention without divulging the manipulations, which came in the third e-mail.

The third e-mail contained the study manipulations and came directly from the SVR Committee member, with an attached signed memo on letterhead, indicating that they were following up on an accusation made by the study's research assistant that the volunteer who graded their session's tests had allegedly done something wrong. Because the participant in question decided to appeal the accusation, SVR policy required formal input from all those who were also present during this session, regardless of their level of involvement. Included in the e-mail was a link to an online survey, which was developed to appear official and standardized, as the SVR supposedly requested this type of feedback often and across various situations. This survey contained our manipulations and included questions to assess participants' perceptions of the fictitious volunteer as a function of the incident.

Subsequent to filling out the SVR survey, participants received a fourth e-mail directly from the researcher. This was the e-mail they had originally anticipated receiving for Phase 2, as it contained a bogus set of questions about taking a test in the presence of others, as well as demographic questions. This e-mail also identified the top five scorers in their session. Finally, after all 10 sessions were completed, a fifth e-mail was sent to all participants that announced the three prize winners for the study and provided a full study debrief.[3]

Manipulations. Eight versions of the SVR survey were developed in accordance with our 2 × 4 between-subjects design. The bulk of each SVR survey was identical; the only differences within the SVR survey were in the allegation and response, which were provided as a function of the formal appeals process. The framing of, and response to, this accusation functioned as the study manipulations.

Intended beneficiary. Similar to Study 1, the trust violation was framed as a transgression that was committed either on behalf of the fictitious volunteer or for another participant in the session. In both conditions, the volunteer was accused of intentionally changing the score of one of the tests the volunteer had graded for the session. More specifically, the research assistant that reported this incident stated "that it was likely intentional, given the volunteer's apparent discomfort when they mentioned that they were double-checking all the scores." In the *self* condition, the accusation stated that "although the volunteer correctly scored other tests, this person changed their own score to be higher than what they actually got." In the *other* condition, the accusation stated that "although the volunteer correctly scored other tests, this person changed the score of another member of the session to be higher than what they actually got."

Violation response. The accusation was then followed by a response from the volunteer, which either apologized for the act in question or combined that apology with a justification. In the

apology condition, the volunteer stated, "I am truly sorry, what the research assistant said about me intentionally changing [my/a session member's] score to be higher than what [I/they] actually got is true." In the apology-plus-justification conditions, the volunteer's response supplemented that apology with one of three distributive justice rationales to explain why the volunteer committed the transgression. Each of these rationales was based on the same underlying logic, namely, that although the difficulty of the questions was supposed to be random, the volunteer (other participant) received twice as many hard questions as everyone else in the session, which seemed unfair:

> When I was doing the scoring, I could see the difficulty levels of all the questions each person received. What I saw just didn't seem fair. Even though the questions were supposed to be random, [I/that student] wound up getting literally twice as many hard questions as everyone else in my session.

In the apology-plus-equity condition, the volunteer continued by stating that "I gave [myself/this person] extra points so that it seemed fairer based on [my/their] obviously high performance on those harder questions." In the apology-plus-equality condition, the volunteer continued by stating that "I gave [myself/this person] extra points so that it seemed fairer based on the fact that [I/they] didn't receive an equal distribution of questions." In the apology-plus-need condition, the volunteer continued by stating that "I gave [myself/this person] extra points so that it seemed fairer so that [I'd/they'd] have a chance to win something [I/they] really need."

Finally, after each of these responses (i.e., after the apology alone, as well as each of the apology-plus-justification responses), the volunteer concluded by stating, "Again, I am really sorry. Please tell me if there's anything I can do to rectify the situation. I hope this makes a difference in the appeals process. If you have any additional questions, please let me know."

Measures. Trusting Intentions ($M = 2.64$, $SD = 1.07$, $\alpha = .83$) and Perceived Fairness ($M = 2.60$, $SD = 1.31$, $\rho = .81$)[4] were assessed with the same items used in Study 1.

Pretests

Pretest #1. To assess the effectiveness of our manipulations without jeopardizing the perceived realism of this study, we first validated our manipulations through an initial pretest. Our initial pilot testing revealed that this separate assessment was particularly necessary for this behavioral study in order to maintain participants' belief that the transgression had been committed, given that participants often became immediately suspicious of the study's authenticity when any sort of manipulation-check question was asked. Eighty-eight participants were recruited from MTurk and were each paid $1 for their involvement. Participants averaged 35.0 years of age ($SD = 12.2$), 6.3 years of work experience ($SD =$

---

[3] During pilot testing, we assessed the frequency and attentiveness with which participants checked their school e-mail in order to determine whether a study involving the receipt of five e-mails over the course of 4 days would be feasible. Overwhelmingly, participants indicated that such communication would be easily digested, as they typically checked their e-mail at least 20 times per day, particularly during the time frame when they are participating in studies for course credit.

[4] The Spearman-Brown statistic is used to calculate reliability when a scale only has two items (Eisinga, Grotenhuis, & Pelzer, 2013).

6.2), and 52% were male. With regard to ethnicity, 75% of these participants were Caucasian, 17% were South Asian, 6% were East Asian, and 2% were African American.

After being randomly assigned to one of the eight conditions, participants were asked to read a description of the incident in the main study and answer eight questions. The first two questions asked whether they recognized the intended beneficiary (self vs. other) as well as the response to the violation (apology plus equity vs. apology plus equality vs. apology plus need vs. apology alone). Eighty-five participants correctly identified the intended beneficiary (97%) and 77 participants correctly identified the response (88%), supporting the effectiveness of the study's manipulations.

The remaining questions from the pretest then assessed whether our operationalizations of these justifications were successful. In this regard, it is critical to note that although each of these justifications could have been operationalized in many different ways, the relevant scientific criteria for evaluating them should not be how one might have personally done so (i.e., a matter of idiosyncratic preference), but rather the more systematic assessment of *whether each operationalization worked as intended and to the same degree* (i.e., so their relative implications could be compared across a level playing field). Thus, we assessed whether these justifications were equivalent in three distinct ways. First, after participants responded to the manipulation-check question related to the response, two items asked the extent to which that response evoked that intended cognition (i.e., equity, equality, or need): (a) "I think the volunteer's action was intended to address that concern," and (b) "I believe the volunteer acted in this way to address that concern" ($\alpha = .89$). Second, two items asked how dire the situation was: (a) "How dire (i.e., terrible) do you think this situation is for the person?" and (b) "How significant do you think this situation is to the person?" ($\alpha = .75$). And finally, two items asked how strong the justifications seemed: (a) "The volunteer's response was strong," and (b) "The volunteer's response was powerful" ($\alpha = .90$). A one-way ANOVA revealed that participants thought that each justification equivalently evoked its intended cognition ($M_{equity} = 5.14$, 95% CI [4.55, 5.73]; $M_{equality} = 5.17$, 95% CI [4.59, 5.75]; $M_{need} = 5.11$, 95% CI [4.78, 5.44]), $F(1, 57) = .01$, $p = .99$, $\eta^2 = .00$, saw the situation as equivalently dire ($M_{equity} = 4.78$, 95% CI [4.02, 5.54]; $M_{equality} = 4.50$, 95% CI [3.91, 5.09]; $M_{need} = 4.59$, 95% CI [4.00, 5.18]), $F(1, 57) = .17$, $p = .84$, $\eta^2 = .00$, and perceived all three justifications to be equally strong ($M_{equity} = 2.79$, 95% CI [2.28, 3.30]; $M_{equality} = 2.91$, 95% CI [2.39, 3.43]; $M_{need} = 3.12$, 95% CI [2.77, 3.47]), $F(1, 57) = .44$, $p = .65$, $\eta^2 = .00$. These results provide clear and consistent support for the adequacy of our manipulations.

Pretest #2. We then conducted a second pretest to verify that trust would indeed be violated as a result of this transgression with a 2 (intended beneficiary: self vs. other) × 2 (time: previolation vs. postviolation) repeated measures design that assessed participants' trusting intentions toward the trustee twice after reading about the study context, once immediately before learning about the transgression and then immediately after learning about the transgression but prior to the trustee's response. Forty-one participants were recruited from MTurk and were each paid $1 for their involvement. Participants averaged 30.32 years of age, 7.0 years of work experience ($SD = 7.2$), and 68% were male. With regard to ethnicity, 76% were Caucasian, 17% were South Asian, and 7% were East Asian. No participants failed the intended beneficiary-

manipulation-check question, but two participants failed the attention-check question and were excluded from the sample.

The repeated measures ANOVA revealed that trusting intentions toward the trustee significantly dropped from before knowing about the transgression to after, $F(1, 39) = 63.89$, $p < .001$, $\eta^2 = .62$, 95% CI [1.23, 2.07]. Moreover, this drop in trust did not significantly differ based on whether the intended beneficiary was the self or another, as evidenced by the interaction failing to reach significance, $F(1, 39) = 0.34$, $p = .56$, $\eta^2 = .01$.

Results

We also conducted three additional validation checks in the main study. First, to provide another assessment of the suitability of this context for investigating trust violations, participants were asked, after reading the allegation, to indicate what their trust had been before, and then after, hearing about the incident. These measures, which were asked prior to hearing the response, thus capture the degree to which participants believed their trust had dropped specifically as a function of hearing about the allegation. A paired t test comparing these pre- ($M = 5.24$, 95% CI [4.94, 5.54]) versus posttrust measures ($M = 2.62$, 95% CI [2.33, 2.91]) revealed a significant drop in participant's trust in the volunteer as a result of the transgression, $t(1,164) = 18.35$, $p < .001$, $d = 1.43$. Moreover, we noted that the posttrust measure, which is effectively the starting point from which participants would receive the trustee's response, did not significantly differ across conditions, as evidenced by the main effects for Intended Beneficiary and Violation Response, as well as the overall interaction not reaching significance ($F[1, 164] = .06$, $p = .82$, $\eta^2 = .05$; $F[1, 164] = 1.22$, $p = .31$, $\eta^2 = .06$; $F[3, 162] = 1.80$, $p = .15$, $\eta^2 = .05$, respectively), ensuring an equivalent starting point for assessing our hypotheses. Finally, the bogus survey (attached to the fourth e-mail) directly asked participants whether they were concerned about anything in the study and ensuing situation with the volunteer, in order to assess whether they had any suspicion about the experience. Only seven participants raised a concern, and we excluded them from the sample.

Confirmatory factor analyses indicated a good fit and supported convergent validity in this study for a two-factor model that included Trusting Intentions and Perceived Fairness ($r = .65$), $\chi^2(12, N = 165) = 19.27$, CFI = .99, NFI = .97, TLI = .98, RMSEA = .061, all item-factor loadings ≥42 ($p < .001$). Moreover, discriminant analyses (Bagozzi & Phillips, 1982) indicated that the hypothesized two-factor model fit the data significantly better than a one-factor model, $\chi^2(13, N = 165) = 47.01$, $\Delta\chi^2(1, N = 165) = 27.74$, $p < .001$. Table 3 reports variable means and standard deviations by condition.

Hypothesis Tests

To assess how justifications would affect the implications of apologies for Trusting Intentions, and how this would depend on the intended beneficiary, we first conducted a 2 (Intended Beneficiary: self vs. other) × 4 (Violation Response: apology plus equity vs. apology plus equality vs. apology plus need vs. apology alone) ANOVA. This analysis revealed a significant main effect for Intended Beneficiary, $F(1, 157) = 7.89$, $p < .01$, $\eta^2 = .05$, 95% CI [.17, .81]. The analysis also revealed a significant main

Table 3
Study 2: Means, Standard Deviations, and CIs by Condition

	M	SD	95% CI	M	SD	95% CI
Justification (plus apology)	Self (*n* = 66)			Other (*n* = 65)		
Trusting intentions	2.27	1.03	[2.02, 2.52]	2.94	1.02	[2.69, 3.19]
Perceived fairness	2.40	1.27	[2.09, 2.71]	2.97	1.32	[2.65, 3.29]
Apology (alone)	Self (*n* = 17)			Other (*n* = 17)		
Trusting intentions	2.92	1.24	[2.33, 3.51]	2.72	0.90	[2.29, 3.15]
Perceived fairness	2.59	1.54	[1.86, 3.32]	1.97	0.74	[1.62, 2.32]
Equity	Self (*n* = 19)			Other (*n* = 20)		
Trusting intentions	2.49	1.00	[2.04, 2.94]	3.28	1.02	[2.83, 3.73]
Perceived fairness	2.84	1.37	[2.22, 3.46]	3.30	1.25	[2.75, 3.85]
Equality	Self (*n* = 24)			Other (*n* = 25)		
Trusting intentions	2.54	1.00	[2.14, 2.94]	2.85	1.08	[2.43, 3.27]
Perceived fairness	2.50	1.14	[2.04, 2.96]	3.30	1.27	[2.80, 3.80]
Need	Self (*n* = 23)			Other (*n* = 20)		
Trusting intentions	1.80	0.97	[1.40, 2.20]	2.70	0.88	[2.31, 3.09]
Perceived fairness	1.93	1.19	[1.44, 2.42]	2.23	1.19	[1.71, 2.75]

Note. CI = Confidence Interval.

effect for Violation Response, $F(1, 157) = 3.27$, $p < .05$, $\eta^2 = .06$. However, these findings were ultimately qualified by a marginally significant Intended Beneficiary × Violation Response interaction, $F(1, 157) = 2.30$, $p = .08$, $\eta^2 = .05$.

Then, to Test Hypothesis 1, which predicted that an apology that has been combined with a justification would result in higher trust (than an apology alone) when the violation was intended to benefit another party (than when the violation was intended to benefit the self), we conducted the same planned contrast of this interaction effect as in Study 1 (Contrast #1). This analysis revealed a significant Intended Beneficiary × Violation Response interaction, $F(1, 157) = 4.88$, $p < .05$, $\eta^2 = .03$. This interaction reveals that although an apology plus justification did not elicit significantly higher Trusting Intentions than an apology alone when the violation was committed to benefit another, $F(1, 78) = 0.70$, $p = .40$, $\eta^2 = .00$, an apology plus justification resulted in significantly lower Trusting Intentions than an apology alone when the violation was committed to benefit the self, $F(1, 79) = 5.05$, $p < .05$, $\eta^2 = .06$ (see Figure 4). These findings support Hypothesis 1.

Next, to evaluate Hypothesis 2, which predicted that an apology combined with an equity-based justification would result in higher trust than an apology combined with a justification based on equality or need, we conducted the same planned contrast for this effect as in Study 1 (Contrast #2). This contrast revealed that the addition of equity-based justifications to an apology resulted in higher Trusting Intentions than the addition of equality- or need-based justifications, $F(1, 157) = 4.57$, $p < .05$, $\eta^2 = .03$, thus supporting Hypothesis 2 (see Figure 5).

Then, to evaluate Hypothesis 3, which predicted that an apology combined with a need-based justification when the intended beneficiary is the self would result in lower trust than an apology combined with an equity- or equality-based justification when the intended beneficiary is the self or an apology combined with equity-, equality-, or need-based justifications when the intended

beneficiary is another, we conducted Contrast #3 (as detailed in Study 1). This analysis revealed that the addition of a need-based justification to an apology when the intended beneficiary was the self led to significantly lower Trusting Intentions than the remaining apology plus justification conditions, $F(1, 162) = 17.45$, $p < .001$, $\eta^2 = .10$, thus supporting Hypothesis 3 (see Figure 5).

Mediation Analyses

Finally, we tested Hypothesis 4's prediction that perceptions of fairness would mediate the Intended Beneficiary × Violation Response interaction's effects on participants' Trusting Intentions by using the same moderated mediation analysis detailed in Study 1 (see Figure 6). The results revealed that Perceived Fairness significantly mediated the effects for Trusting Intentions, as evidenced by the fact that the 95% CI did not include zero [−.31, −.04]. This supports Hypothesis 4 (see Table 4).
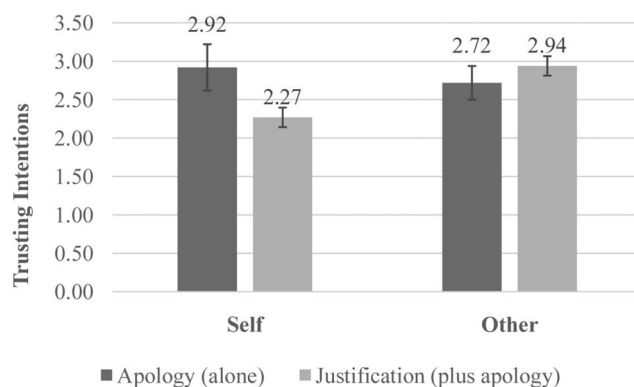

Figure 4. Study 2: Effects of an apology versus justification as a function of the intended beneficiary.
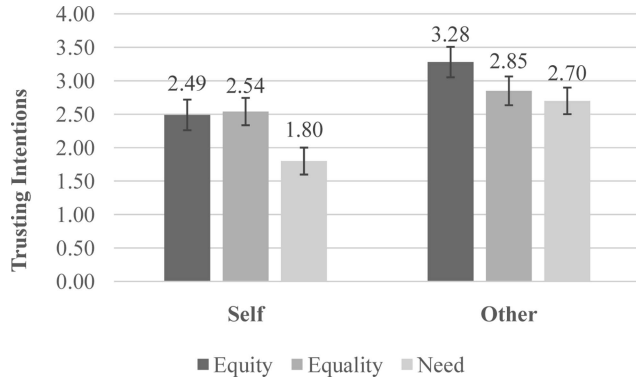
Figure 5. Study 2: Effects of equity, equality, and need as a function of the intended beneficiary.

General Discussion

The purpose of this article was to investigate how three of the most common justifications from the distributive justice literature (equity, equality, and need) would affect the implications of apologies for trust after its violation, and how those relationships would depend on the intended beneficiary. To explore this question, we conducted two experiments, including a new type of study that supplemented the rigor of traditional laboratory methods with far greater realism, and found highly consistent results across these studies despite their methodological differences. In this regard, it is critical to understand that these findings do not center on overall main effect differences among these responses (e.g., Hypothesis 2's attempt simply to extend past findings to the trust violation context), but rather on how the relative implications of these responses *can change as a function of the intended beneficiary* (i.e., Hypotheses 1 and 3's predicted interactions). This is because past research has made clear that the results of many efforts to address trust violations can depend quite substantially on the specific nature of the situation (e.g., Dirks et al., 2011; Ferrin et al., 2007; Kim et al., 2009). Supporting this claim, the present findings provide strong support for our prediction that the combination of an apology with a justification would result in higher trust (than an apology alone) when the violation was intended to benefit another party (than when the violation was intended to benefit oneself). Finer-grained analyses also reveal that the addition of equity-based justifications to an apology elicited higher trust than the addition of equality- or need-based justifications on the whole, and that the addition of need-based justifications was particularly ineffective when the violation was intended to benefit the self. Finally, me-

diation analyses supported the prediction that these effects would be driven by perceptions of fairness.

Theoretical Implications

Although recent theory has proposed that trust can be influenced by altering the perceived nature of the violation (Kim et al., 2009), the extent to which justifications, in particular, could facilitate that goal has remained empirically unexplored. This inquiry not only addresses this limitation by demonstrating that the addition of justifications can elicit higher trust than apologies on their own, but also extends theory regarding this relationship by revealing how the relative benefits of combining justifications with apologies versus providing apologies alone can depend on the intended beneficiary. Thus, whereas past research on trust violations has examined contexts in which the transgression was committed either for oneself or others without considering whether the person(s) to be affected matters, the present findings suggest that this choice of target is a critical factor that can dramatically influence the level of trust achieved after its violation. By doing so, the findings ultimately challenge the longstanding and widely held assumption by research on both justice and trust that justifications would be useful and beneficial, by revealing that adding such justifications may actually do little for violations intended to benefit the self, and indeed prove to be no better or worse than what had consistently been found to be the least effective response one can offer for an integrity-based trust violation (i.e., an apology).

This finding, in turn, contributes to the distributive justice literature by highlighting how the implications of justifications can differ when they are used before versus after a violation has occurred. Whereas previolation considerations of distributive justice tend to be more systemic, regarding how allocations should occur across all members of a group, trust violations tend to be more targeted (i.e., they are more likely to concern efforts to benefit or harm a subset of individuals, irrespective of the implications for others). This consideration of the target can, thus, make both the use and implications of postviolation justifications more complicated than previolation justifications, and require more theoretical and empirical attention than they have so far received.

Even further, the fact that need-based justifications were particularly ineffective when the violation was intended to benefit the self highlights the importance of considering how people interpret these justifications more carefully. Indeed, the finding suggests that even though people are likely to be in the best position to assess their own need and feel greater pressure to address it when that need is great, this is also precisely when those evaluating
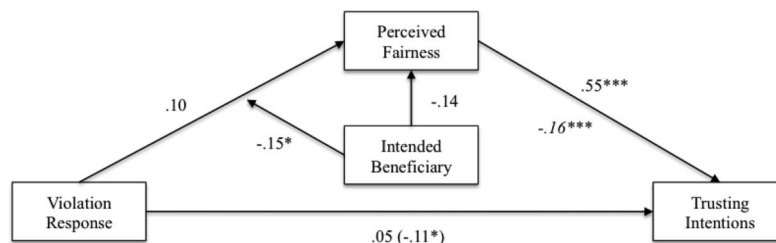

Figure 6. Study 2: Test of moderated mediation.

Table 4
Study 2: Results of Moderated Mediation

	B	SE	t	p	95% CI
Direct effects on fairness					
Violation response	0.10	0.06	1.65	0.101	[−.02, .22]
Intended beneficiary	−0.14	0.10	−1.36	0.176	[−.33, .06]
Violation Response × Intended Beneficiary	−0.15	0.06	−2.4	0.017	[−.26, −.03]
Direct effects on trusting intentions					
Violation response	0.05	0.05	1.05	0.294	[−.16, .05]
Perceived fairness	0.55	0.05	11.38	0.000	[.45, .65]
Indirect effect on trusting intentions					
Violation Response × Intended Beneficiary → Perceived Fairness	−0.16	0.06	a	a	[−.31, −.04]

a. Statistic not applicable when assessing indirect effects and reporting the index of moderated mediation.

trustworthiness would consider such justifications less meaningful. This raises the potential for an ironic conundrum where the very cases in which higher trust might be more warranted after a violation has occurred are the same cases in which it is less likely to be observed.

These issues also highlight the importance of considering how the level of trust after its violation may depend not only on the ramifications of the transgression and subsequent response for the trustor as an individual, but also on what these factors would imply for others. Thus far, virtually all past research on trust violations has adopted a bilateral perspective, by focusing on the interactions between just one trustor and one trustee (see Kim et al., 2009, for a review). The present findings, however, illustrate how broadening this view can allow us to examine an even wider array of methods for addressing such incidents and account for additional mechanisms (e.g., arising from considerations of justice and the intended beneficiary) that can help inform how trust might subsequently be influenced. By doing so, this work suggests that addressing trust violations can be considered a matter not only of addressing concerns about a given one-on-one relationship but also of addressing concerns about the broader collective.

Finally, the fact that the relative benefits of combining justifications with apologies versus providing apologies alone for trust were mediated by perceived fairness supports the present effort to integrate the distributive justice and trust literatures. Although fairness is a central consideration in the distributive justice literature, and this perception has been recognized to be an important determinant of trust, finer-grained studies of distributive justice and trust have largely developed in isolation. However, given that each literature has considered issues that have been largely neglected by the other, as the present inquiry has sought to reveal, their underlying theoretical relationship may allow greater insights to be gained through their combination.

For example, in the organizational justice literature, Folger and Cropanzano's (1998) fairness theory proposes that people gauge the fairness of a situation by evaluating whether an incident (a) caused harm, (b) was under the violator's control, and (c) violated some moral or ethical normative standard. Thus, we can observe that the present inquiry's focus on integrity-based violations met all three criteria (which they label "would," "could," and "should" counterfactuals, respectively), and may be further understood as an attempt by the violator to alter the "should" counterfactual by referring to a competing normative standard (i.e., the maintenance of distributive

justice) that could counteract the perception that the incident had been unfair. In this way, the organizational justice literature offers an alternative way of interpreting our effects that may have the potential to offer insights beyond what the trust repair literature has considered. Yet it is also the case that our research can help advance the organizational justice literature (a) by highlighting the challenges that efforts to alter initial unfairness perceptions can pose, (b) by considering how different attempts to alter "should" counterfactuals can differ markedly in their effectiveness, and (c) by moving beyond this literature's past efforts to consider what might affect these fairness perceptions in an additive sense (e.g., Greenberg, 1994, 2006; Skarlicki & Folger, 1997) to reveal how their relative importance can differ depending on the situation.

Practical Implications, Limitations, and Future Directions

These results can also provide important practical implications for those seeking to address trust violations. In light of the fact that past research has underscored the difficulty of addressing integrity-based trust violations (Ferrin et al., 2007; Kim et al., 2004, 2006, 2013), the present findings offer a solution. That is, rather than simply apologize or take the unethical and risky path of trying to deny culpability when one is actually culpable, trustees may instead consider combining an apology with a justification when their culpability is clear. However, the findings also underscore how the benefits of doing so can depend on the intended target of the violation, as well as the type of justification used, and the consequent importance of considering when these justifications are more or less likely to help. Of particular note in this regard is the fact that even though one might be in the best position to gauge need, and have less liberty to ignore it when that need concerns oneself than another, it may actually be better not to offer such justifications after a violation and instead simply apologize.

Such conclusions should also be cautioned, however, by several limitations. First, although we found support for our predictions in two different contexts, there are at least two reasons why these findings may not necessarily generalize to other kinds of settings. More specifically, the distributive justice literature has long recognized that, although equity may be considered more important than equality or need in performance-oriented relationships, equality or need may be considered more important than equity when performance is not the overriding goal (Conlon et al., 2004;

Deutsch, 1975; Mannix et al., 1995). Even further, one might observe that the participants in this study were primarily from the United States. Yet cross-cultural research suggests that Indians, for example, are more likely to distribute resources on the basis of need and less on the basis of equity or equality than Americans (Murphy-Berman, Berman, Singh, Pachauri, & Kumar, 1984). These kinds of differences, therefore, suggest that support for our predictions is likely to depend, at least to some degree, on the extent to which performance has been prioritized in that setting as well as the culture in which these assessments are made.

Second, despite Study 2's efforts to bolster psychological realism and external validity, its focus was still restricted to a violator that participants did not really know, and it also lacked a meaningful behavioral measure of trust. This focus on newly formed relationships does not imply that trust violations failed to occur, given that people have been observed to exhibit surprisingly high levels of initial trust in others even without a history of interaction (Harmon, Kim, & Mayer, 2014; McKnight, Cummings, & Chervany, 1998), and both this and other studies reveal that such high initial trust can be significantly lowered by an alleged transgression (Dirks et al., 2011; Ferrin et al., 2007; Kim et al., 2004). Nevertheless, it would be useful to extend this research in the future by investigating the repair of trust in longer-term relationships and by examining the implications of such responses for parties' actual behaviors as well.

Third, it might also be useful to explore the implications of operationalizing these justifications and intended beneficiaries in different ways. For example, the equity-, equality-, and need-based justifications might have all been made more or less effective by altering how each justification had been explained; the magnitude of the initial inequity, inequality, or need that the violation had been intended to address; or whether the justification had been accompanied by an apology. Although each of these possibilities would be limited to influencing the overall effectiveness of these justifications, rather than the predicted interactions that represent the focus of this inquiry, it would be helpful to account for such effects. Likewise, one might investigate how our effects might differ if the violator had actually intended to benefit the party assessing the violator's trustworthiness, rather than a third party, or if the violation had been intended to benefit both the self and another, rather than the self alone. Presumably, efforts to justify the former alternative would prove more effective than the latter in each comparison, and even more so if such outcomes seemed more deserved, but this would require empirical validation. Even further, it would be useful for future research to assess parties' interpretations of and reactions to these justifications in greater detail (e.g., by asking participants to report their own interpretations of these justifications, rather than answer multiple choice questions about them, and by doing so in the actual study rather than via pretests) to provide greater insight into the specific kinds of mechanisms through which each might operate.

Finally, although the present study's focus on equity, equality, and need was guided by the significant attention they have received in the distributive justice literature, it is important to recognize that other justifications may be used for trust violations as well. These include justifications based on group membership (Kramer, 1991), duty (Hosmer, 1995), or emotions such as anger and revenge (Bies & Tripp, 1996), to name just a few examples. Thus, it would be useful to expand research on addressing trust violations to investigate the effects of other justifications to develop a better sense of their relative implications.

Conclusion

Despite marked growth in research on trust violations, the literature still offers little insight on how trust might be influenced after an integrity-based violation, particularly when one is actually culpable. The present effort begins to address this limitation, not only by demonstrating the potential benefits of adding justifications to an apology when one is guilty, but also by revealing how these potential benefits can depend on the intended beneficiary and type of justification used. Thus, even though justifications may not always be beneficial, the present insights into when this might be the case should provide at least some important guidance regarding such incidents.

References

Adams, J. S. (1965). Inequity in social exchange. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 267–299). New York, NY: Academic Press.

Austin, W. (1980). Friendship and fairness: Effects of type of relationship and task performance on choice of distribution rules. *Personality and Social Psychology Bulletin, 6,* 402–408. doi:10.1177/014616728063011

Babcock, L., Loewenstein, G., Issacharoff, S., & Camerer, C. (1995). Biased judgements of fairness in bargaining. *American Economic Review, 85,* 1337–1342.

Bagozzi, R. P., & Phillips, L. W. (1982). Representing and testing organizational theories: A holistic construal. *Administrative Science Quarterly, 27,* 459–489.

Bell, B. E., & Loftus, E. F. (1989). Trivial persuasion in the courtroom: The power of (a few) minor details. *Journal of Personality and Social Psychology, 56,* 669–679. doi:10.1037/0022-3514.56.5.669

Bies, R. J., & Tripp, T. M. (1996). Beyond distrust: Getting even and the need for revenge. In R. M. Kramer & T. R. Tyler (Eds.), *Trust in organizations: Frontiers of theory and research* (pp. 246–260). Thousand Oaks, CA: Sage. doi:10.4135/9781452243610.n12

Bottom, W. P., Gibson, K., Daniels, S., & Murnighan, J. K. (2002). When talk is not cheap: Substantive penance and expressions of intent in rebuilding cooperation. *Organization Science, 13,* 497–513. doi:10.1287/orsc.13.5.497.7816

Buckless, F. A., & Ravenscroft, S. P. (1990). Contrast coding: A refinement of ANOVA in behavioral analysis. *Accounting Review, 65,* 933–945.

Cohen-Charash, Y., & Spector, P. E. (2001). The role of justice in organizations: A meta-analysis. *Organizational Behavior and Human Decision Processes, 86,* 278–321. doi:10.1006/obhd.2001.2958

Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O. L. H., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology, 86,* 425–445. doi:10.1037/0021-9010.86.3.425

Conlon, D. E., Porter, C. O. L. H., & Parks, J. M. (2004). The fairness of decision rules. *Journal of Management, 30,* 329–349. doi:10.1016/j.jm.2003.04.001

Conway, N., & Briner, R. B. (2002). A daily diary study of affective responses to psychological contract breach and exceeded promises. *Journal of Organizational Behavior, 23,* 287–302. doi:10.1002/job.139

Desmet, P., De Cremer, D., & Van Dijk, E. (2011). In money we trust? The use of financial compensations to repair trust in the aftermath of distributive harm. *Organizational Behavior and Human Decision Processes, 114,* 75–86. doi:10.1016/j.obhdp.2010.10.006

Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social Issues, 31,* 137–149. doi:10.1111/j.1540-4560.1975.tb01000.x

Deutsch, M. (1985). *Distributive justice: A social-psychological perspective.* New Haven, CT: Yale University Press.

Dirks, K. T., Kim, P. H., Ferrin, D. L., & Cooper, C. D. (2011). Understanding the effects of substantive responses on trust following a transgression. *Organizational Behavior and Human Decision Processes, 114,* 87–103. doi:10.1016/j.obhdp.2010.10.003

Eisinga, R., Grotenhuis, M. T., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health, 58,* 637–642. doi:10.1007/s00038-012-0416-3

Fehr, E., & Fischbacher, U. (2004). Third party punishment and social norms. *Evolution and Human Behavior, 25,* 63–87. doi:10.1016/S1090-5138(04)00005-4

Ferrin, D. L., Kim, P. H., Cooper, C. D., & Dirks, K. T. (2007). Silence speaks volumes: The effectiveness of reticence in comparison to apology and denial for responding to integrity- and competence-based trust violations. *Journal of Applied Psychology, 92,* 893–908. doi:10.1037/0021-9010.92.4.893

Greenberg, J. (1994). Using socially fair treatment to promote acceptance of a work site smoking ban. *Journal of Applied Psychology, 79,* 288–297.

Greenberg, J. (2006). Losing sleep over organizational injustice: Attenuating insomniac reactions to underpayment inequity with supervisory training in interactional justice. *Journal of Applied Psychology, 91,* 58–69.

Folger, R., & Cropanzano, R. (1998). *Organizational justice and human resource management.* Thousand Oaks, CA: Sage.

Harmon, D. J., Kim, P. H., & Mayer, K. J. (2014). Breaking the letter vs. spirit of the law: How the interpretation of contract violations affects trust and the management of relationships. *Strategic Management Journal.* Advance online publication. doi:10.1002/smj.2231

Hosmer, L. T. (1995). Trust: The connecting link between organizational theory and philosophical ethics. *The Academy of Management Review, 20,* 379–403.

Kim, P. H., Cooper, C. D., Dirks, K. T., & Ferrin, D. L. (2013). Repairing trust with individuals vs. groups. *Organizational Behavior and Human Decision Processes, 120,* 1–14. doi:10.1016/j.obhdp.2012.08.004

Kim, P. H., Dirks, K. T., & Cooper, C. D. (2009). The repair of trust: A dynamic bilateral perspective and multilevel conceptualization. *The Academy of Management Review, 34,* 401–422. doi:10.5465/AMR.2009.40631887

Kim, P. H., Dirks, K. T., Cooper, C. D., & Ferrin, D. L. (2006). When more blame is better than less: The implications of internal vs. external attributions for the repair of trust after a competence- vs. integrity-based trust violation. *Organizational Behavior and Human Decision Processes, 99,* 49–65. doi:10.1016/j.obhdp.2005.07.002

Kim, P. H., Ferrin, D. L., Cooper, C. D., & Dirks, K. T. (2004). Removing the shadow of suspicion: The effects of apology vs. denial for repairing ability- vs. integrity-based trust violations. *Journal of Applied Psychology, 89,* 104–118. doi:10.1037/0021-9010.89.1.104

Kramer, R. M. (1991). Intergroup relations and organizational dilemmas: The role of categorization processes. *Research in Organizational Behavior, 13,* 191–228.

Kramer, R. M., & Lewicki, R. J. (2010). Repairing and enhancing trust: Approaches to reducing organizational trust deficits. *The Academy of Management Annals, 4,* 245–277. doi:10.1080/19416520.2010.487403

Lazarus, R. S., & Cohen-Charash, Y. (2001). Discrete emotions in organizational life. In R. Payne & C. Cooper (Eds.), *Emotions at work* (pp. 45–81). Hoboken, NJ: Wiley.

Lount, R. B., Jr., Zhong, C.-B., Sivanathan, N., & Murnighan, J. K. (2008). Getting off on the wrong foot: The timing of a breach and the restoration of trust. *Personality and Social Psychology Bulletin, 34,* 1601–1612. doi:10.1177/0146167208324512

Maddux, W. W., Kim, P. H., Okumura, T., & Brett, J. M. (2011). Cultural differences in the function and meaning of apologies. *International Negotiation, 16,* 405–425. doi:10.1163/157180611X592932

Mannix, E. A., Neal, M. A., & Northcraft, G. B. (1995). Equity, equality, or need? The effects of organizational culture on the allocation of benefits and burdens. *Organizational Behavior and Human Decision Processes, 63,* 276–286. doi:10.1006/obhd.1995.1079

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review, 20,* 709–734.

McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial trust formation in new organizational relationships. *The Academy of Management Review, 23,* 473–490.

Messick, D. M., & Sentis, K. P. (1979). Fairness and preference. *Journal of Experimental Social Psychology, 15,* 418–434. doi:10.1016/0022-1031(79)90047-7

Murphy-Berman, V., Berman, J. J., Singh, P., Pachauri, A., & Kumar, P. (1984). Factors affecting allocation to needy and meritorious recipients: A cross-cultural comparison. *Journal of Personality and Social Psychology, 46,* 1267–1272. doi:10.1037/0022-3514.46.6.1267

Nelissen, R. M. A., & Zeelenberg, M. (2009). Moral emotions as determinants of third-party punishment: Anger, guilt, and the functions of altruistic sanctions. *Judgment and Decision Making, 4,* 543–553.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40,* 879–891. doi:10.3758/BRM.40.3.879

Reis, H., & Gruzen, J. (1976). On mediating equity, equality, and self-interest: The role of self presentation in social exchange. *Journal of Experimental Social Psychology, 12,* 487–503. doi:10.1016/0022-1031(76)90079-2

Robinson, S. L., & Rousseau, D. M. (1994). Violating the psychological contract: Not the exception but the norm. *Journal of Organizational Behavior, 15,* 245–259. doi:10.1002/job.4030150306

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. doi:10.5465/AMR.1998.926617

Scher, S. J., & Darley, J. M. (1997). How effective are the things people say to apologize? Effects of the realization of the apology speech act. *Journal of Psycholinguistic Research, 26,* 127–140. doi:10.1023/A:1025068306386

Schweitzer, M. E., Hershey, J. C., & Bradlow, E. T. (2006). Promises and lies: Restoring violated trust. *Organizational Behavior and Human Decision Processes, 101,* 1–19. doi:10.1016/j.obhdp.2006.05.005

Scott, M. B., & Lyman, S. M. (1968). Accounts. *American Sociological Review, 33,* 46–62. doi:10.2307/2092239

Shaw, J. C., Wild, E., & Colquitt, J. A. (2003). To justify or excuse? A meta-analytic review of the effects of explanations. *Journal of Applied Psychology, 88,* 444–458. doi:10.1037/0021-9010.88.3.444

Skarlicki, D. P., & Folger, R. (1997). Retaliation in the workplace: The roles of distributive, procedural, and interactional justice. *Journal of Applied Psychology, 82,* 434–443.

Stouten, J., De Cremer, D., & Van Dijk, E. (2006). Violating equality in social dilemmas: Emotional and retributive reactions as a function of trust, attribution, and honesty. *Personality and Social Psychology Bulletin, 32,* 894–906. doi:10.1177/0146167206287538

Tomlinson, E. C., Dineen, B. R., & Lewicki, R. J. (2004). The road to reconciliation: Antecedents of victim willingness to reconcile following a broken promise. *Journal of Management, 30,* 165–187. doi:10.1016/j.jm.2003.01.003