Journal of Experimental Psychology: Applied

I Am (Not) Sorry: Interpersonal Effects of Neutralizations After a Transgression

Bastiaan T. Rutjens, Coen A. Ackers, and Gerben A. van Kleef

Online First Publication, August 17, 2023. https://dx.doi.org/10.1037/xap0000483

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: Applied

ISSN: 1076-898X

https://doi.org/10.1037/xap0000483

I Am (Not) Sorry: Interpersonal Effects of Neutralizations After a Transgression

Bastiaan T. Rutjens, Coen A. Ackers, and Gerben A. van Kleef
Department of Psychology, University of Amsterdam

After a transgression, people often use neutralizations to account for their behavior, for instance, by apologizing or offering a justification. Previous research has mostly centered around the intrapersonal effects of neutralizations on actors. Consequently, we know very little of the *interpersonal* effects of neutralizations on observers' perceptions and judgments. Our overarching hypothesis is that neutralizations that contain an acknowledgment of wrongdoing (i.e., apologies and excuses) lead to more favorable perceptions of the transgressor and the transgression than neutralizations that do not (i.e., justifications). We report three studies ($N = 800$) to investigate the relationship between the type of neutralization used and observers' perceptions of actors and their behaviors. Our findings show that actor and behavior are evaluated differently depending on whether the neutralization used is an apology, an excuse, a consequentialist justification, or a deontological justification. Overall, justifications led to more negative evaluations (especially when invoking deontological reasoning), while apologies and excuses fostered more positive evaluations. We discuss the implications of these findings for understanding the social dynamics of norm violations and the social and legal implications for enforcing norm abidance.

Public Significance Statement
People use various types of neutralizations to account for a transgression they have committed, and these can range from apologies to justifications. The current research focuses on the effects of the type of neutralization used on how other people subsequently evaluate the transgressor. These insights increase understanding of the social consequences of using various types of neutralizations.

Keywords: norm violation, excuses, justifications, moral disengagement, neutralizations

Suppose an ordinary individual named Jack intentionally did something he knows most others would consider morally wrong, such as removing a 100-dollar note from a found wallet before returning it. In case he gets caught, Jack might wonder, would the explanation he might give for this behavior affect other people's judgments of him and his behavior? What sort of account would work best to reduce the risk of condemnation? What if Jack apologized or offered an excuse? What if he argued that the end justified the means or that he had the right to do what he did? Would observers think differently of Jack and his behavior depending on the kind of account he offered? Would they consider him a more or less moral person? Would they condemn his behavior more or less harshly and advocate a more or less severe penalty?

These are important questions, not just for occasional wrongdoers like Jack but for managing moral transgressions and sustaining societal functioning more broadly (Van Kleef et al., 2019). Individuals are motivated to avoid disapproval, especially regarding moral behavior. Not only is moral condemnation a precursor to exclusion and other social penalties, but it can also do damage to one's self-concept (Leach et al., 2007; Van der Lee et al., 2017). From a societal and legal point of view, it is also important to understand whether and how judgments of moral transgressors are shaped by actors' accounts—or "neutralizations"—of their behavior. The chief objective of the current research is to investigate the effects of different types of neutralizations used by actors on observers' judgments of actors and their acts.

Bastiaan T. Rutjens https://orcid.org/0000-0003-3163-4156

The authors wish to thank Tisa Bertlich, Julius März, and Natalia Zarzeczna for their assistance with data analysis. The authors have no known conflicts of interest to disclose.

Bastiaan T. Rutjens played a lead role in writing–review and editing and an equal role in conceptualization, data curation, formal analysis, investigation, methodology, project administration, supervision, and writing–original draft. Coen A. Ackers played a lead role in writing–original draft and an equal role in conceptualization, formal analysis, investigation, and methodology. Gerben A. van Kleef played a supporting role in writing–review and editing and an equal role in conceptualization, formal analysis, investigation, methodology, project administration, supervision, and writing–original draft.

Correspondence concerning this article should be addressed to Bastiaan T. Rutjens or Gerben A. van Kleef, Department of Psychology, University of Amsterdam, Nieuwe Achtergracht 129, Postbus 15900, Amsterdam 1001 NK, The Netherlands. Email: b.t.rutjens@uva.nl or g.a.vankleef@uva.nl

Neutralizations and What They Do

We define neutralizations as statements of opinion regarding one's behavior or attitude in reference to a rule or standard, the general validity of which is not itself contested.[1] Such statements can range from a full admission of guilt to a verbal plea for the granting of a temporary or personal, full or partial, exception to the otherwise valid rule. Guided by this definition, we invoke three key criteria to delineate four fundamentally different types of neutralizations, which are described in detail below and summarized in Table 1: whether the wrongness of the act is admitted, whether responsibility for the act is acknowledged, and whether the validity of the neutralization is made dependent on the consequences of the act.

Neutralizations have been examined with regard to a wide variety of troublesome behavior, including criminality in general (Cechaviciute & Kenny, 2007; Van Baak et al., 2018), aggression (Agnew & Peters, 1986), genocide (Bryant et al., 2017), sex trafficking (Copley, 2014), academic cheating (Haines et al., 1986), unethical work behavior (Gauthier, 2001; Hollinger, 1991; Robinson & Kraatz, 1998), and refusing to recycle. Despite this large body of research, three aspects of neutralization use have received little attention.

First, and most importantly, although previous research efforts have been directed at understanding the *intrapersonal* effects of neutralizations—that is, the impact neutralizations have on actors themselves—little is known about the *interpersonal* effects of neutralizations—that is, their impact on observers' judgments. Previous work on third-party evaluations has primarily looked at how people expressing certain moral judgments about situations or acts are evaluated (e.g., Everett et al., 2016), but not at how neutralization use is evaluated. This is remarkable because neutralizations have long been recognized as serving the purpose of influencing observers' judgments. According to Klenowski et al. (2011), "the accounts given must be believable to those offering them and, more importantly, be honored by the social audience" (p. 47). Thus, the choice of neutralization type depends both on its presumed effectiveness to preserve a positive self-image and on its effectiveness in averting social condemnation. In other words, they are means of impression management. Previous research has investigated the interpersonal effects of apologies (e.g., Ohbuchi et al., 1989; Van Kleef & De Dreu, 2010) but paid scant attention to other types of neutralizations. To more fully understand the social dynamics of norm violations and the implications for enforcing norm abidance, understanding the *interpersonal consequences* of different types of neutralizations is crucial.

Second, the effectiveness of different types of neutralizations relative to each other has seldom been explored. As a result, the consequences of using different neutralization techniques for impression management remain poorly understood. For society at large, the interpersonal effects of various types of neutralizations are just as important as their intrapersonal effects. To the degree that people condone others' norm infractions because of the neutralizations they employ, the likelihood of future recurrence of similar infractions may increase, potentially resulting in a more general erosion of norms in the long run (Van Kleef et al., 2015).

Third, previous work has not systematically classified neutralizations. The dimensional approach to neutralizations that we utilize in the current work helps cluster various forms of neutralization in four fundamentally different categories based on theoretically meaningful underlying dimensions (see below). This not only enables a more systematic test of their effectiveness but also allows for broader conclusions than were allowed by previous, more idiosyncratic, approaches that lacked such theoretical grounding.

To test the interpersonal effects of neutralization use, the current work utilizes measures from the social evaluation literature, in particular the stereotype content model (Fiske et al., 2002) and the behavioral regulation model (Leach et al., 2007). Whereas the former has been used to understand perceptions of various social groups across two dimensions—warmth and competence—the latter breaks down warmth into two dimensions: morality and sociability. Evaluations of morality and sociability have been shown to be perceived as diagnostic (i.e., relatively stable across contexts; e.g., Van der Lee et al., 2017), which makes them particularly suitable to assess the effects of using various neutralizations following a moral transgression.

Effects of Different Types of Neutralizations

Based on previous theorizing and empirical work that examined links between neutralizations and behaviors—described above—we develop a differentiation of neutralizations based on three theoretical dimensions on which neutralizations can vary (see Table 1). Specifically, we distinguish among neutralizations based on whether (a) the actor admits that the act was wrong, (b) the actor acknowledges full causal responsibility for the act, and (c) the neutralization depends for its purported validity on the consequences of the act.

These three criteria delineate four types of neutralizations that have been documented previously (please see Footnote 1; also see Cottee, 2010; Schönbach, 2010; Scott & Lyman, 1968; Sykes & Matza, 1957): apologizing, excuse-making, consequentialist justifications, and deontological justifications. As shown in Table 1,

[1] We use the term "neutralization" instead of alternatives such as, in the chronological order they were coined, "vocabularies of motives" (Mills, 1940), "techniques of neutralization" (Sykes & Matza, 1957), "accounts" (Scott & Lyman, 1968), or "mechanisms of moral disengagement" (Bandura, 1990; Moore, 2015). These different terms have meanings that largely or fully overlap. Scott and Lyman (1968) referred to the work of both Mills (1940) and Sykes and Matza (1957), acknowledging the similarity of the terms preferred by these authors. Bandura and colleagues recognized that neutralization theory as proposed by Sykes and Matza (1957) "includes some of the mechanisms of moral disengagement" (Osofsky et al., 2005, p. 373). Akers (2009) declared mechanisms of moral disengagement to be "virtually a clone" (p. 86) of neutralization techniques, while Ribeaud and Eisner (2010) concluded that neutralization techniques and moral disengagement modes "appear to be broadly congruent" (p. 302). Thus, neutralizations and moral disengagement mechanisms are largely overlapping concepts.

Although our definition captures most of the neutralization techniques so far identified, it does not include the use of smokescreens or clearly articulated denials of the behavior or its consequences, such as "denial of injury" (Sykes & Matza, 1957), "euphemistic labeling" (Bandura, 1990), "minimizing, ignoring or misconstruing the consequences" (ibid.), and "palliative comparison" (ibid.), neither does it include deliberate distractions that fail to address the behavior in question, like "refocusing attention" (Ashforth & Anand, 2003), condemnation of condemners (Sykes & Matza, 1957), "avoidance" (Metts & Cupach, 1989), and "postponement" (Cromwell & Thurman, 2003). The reason for excluding these techniques is that neutralization as we define it captures how the actor relates to the transgression, similar to moral disengagement theory (Moore, 2015). That is, if the statement regarding one's behavior is meant to—for example—distract the evaluator or deny the behavior to have taken place in the first place, it does no longer target the transgression, and therefore it is not a neutralization of that transgression.

Table 1
The Fundamental Differences of Apologizing, Excuse-Making, Consequentialist Justification, and Deontological Justification

Type of neutralization	Wrongness admitted	Responsibility acknowledged	Consequence-considered
Apologizing	Yes	Yes	No
Excuse-making	Yes	No	No
Consequentialist justification	No	Yes	Yes
Deontological justification	No	Yes	No

apologies and excuses are neutralizations in which wrongness is admitted. What differentiates apologies and excuses is that by offering an apology, one acknowledges full responsibility for having broken a normative rule, whereas by offering an excuse, one denies full responsibility. By using an excuse, the neutralizing actor declares implicitly that he or she would not have committed the act had it not been for the occurrence of the excuse.

Unlike apologies and excuses, consequentialist and deontological justifications entail no admission of wrongdoing. Consequentialist justifications refer to the effects of the neutralized behavior, argue that the "end justifies the means," and thus make their validity dependent on a per sum beneficial outcome. They appeal to a greater good that can only be secured by the act and which would benefit third parties or even its possible victims. A consequentialist justification is the outcome of weighing the net results of the act in question against the net results of an alternative act. Deontological justifications lack such dependence on the consequences and instead appeal to the actor having a right to conduct the behavior, or to a victim lacking a right to contest that behavior. As such, deontological justifications are supposedly valid irrespective of the consequences of the act.

Current Research and Hypotheses

As laid out above, the three main contributions to the literature on neutralizations that we envisioned with the current work lie in its focus on their interpersonal effects, on their relative effectiveness, and in the systematic classification of neutralization types that we utilized. In developing our hypotheses, we took guidance from the three criteria shown in Table 1. The first criterion distinguishes between neutralizations that contain an acknowledgment of wrongdoing and those that do not. While offerings of apologies and excuses confirm the full validity of the broken norm, justifications contest the applicability of the broken norm for the specific act that is referred to (Cottee, 2010). Therefore, justifications should be considered more of a threat to the norm's continued existence than the other two neutralizations that reaffirm adherence to the norm without qualifications. Consequently, neutralizations that contain an admission of wrongdoing and an implied intention to respect the norm in the future may be perceived as more acceptable or convincing than those that do not.

In addition, it seems likely that the use of neutralizations that are considered less acceptable or convincing by observers indicates that the neutralizing actor is further removed from mainstream morality. This in turn leads us to assume that the recipients of a neutralization that fails to convince will also hold the person using that neutralization personally in lower regard (at least in terms of their morality). We therefore propose that observers have a more positive view of actors who admit wrongdoing than of those who do not. To test this, we borrow from studies in the field of person perception. Fiske et al. (2002) identified warmth and competence as two fundamental dimensions underlying person perception, with the dimension of warmth in turn consisting of two distinct subdimensions: sociability and morality (Leach et al., 2007; see also Brambilla et al., 2011; Goodwin et al., 2014). We hypothesize that observers consider actors using an apology or an excuse as possessing a higher level of morality (Hypothesis 1a) and sociability (Hypothesis 1b) than actors who use a justification. Given that we focus on norm violations (i.e., transgressions), we do not expect to find any differences for perceptions of competence (also see Footnote 3). Further, we predict that observers condemn behavior by actors who admit wrongdoing less severely than behavior accompanied with a justification in terms of their evaluation of the behavior (Hypothesis 1c) and their preferred punishment of the actor (Hypothesis 1d).

Besides testing these formal hypotheses regarding differences between apologies and excuses on the one hand and consequentialist and deontological justifications on the other hand, we set out to examine potential differences within each of these two clusters. First, distinguishing between offering an apology and making excuses, it could be reasoned that offering an excuse—and thereby denying full causal responsibility for one's behavior—presupposes taking less responsibility or commitment to avoid future transgressions than offering an apology. Accordingly, observers might make harsher judgments when a perpetrator offers an excuse than when he or she offers an apology. Second, it is conceivable that a consequentialist justification contrasts less strongly with the violated norm than a deontological justification. The validity of the former depends on the occurrence of particular circumstances that produce a greater benefit of transgressive behavior compared to norm adherence, whereas the validity of the latter does not depend on incidental circumstances. Therefore, future transgressions would presumably be less likely if the actor uses a conditional, consequentialist justification than if he or she uses an unconditional, deontological justification. In other words, a justification that depends on the consequences may be judged less harshly than a justification that does not. We examined these possibilities in an exploratory fashion in a pilot study and followed up on them in two additional studies as appropriate.

We conducted three studies (a pilot and two main studies), in which we presented participants with scenarios describing mild norm infractions. Ethics approval was obtained at the University of Amsterdam, Psychology Research Institute (Project 2016-SP-7517).

All scenarios can be found in Appendices A and B. Constituting six different conditions[2] to which each participant was assigned, the scenario was followed by an account given by the protagonist representing one of the four forms of neutralizations described above, no account at all (dishonest control condition), or a scenario without norm infractions (honest control condition). The control conditions allowed for a baseline comparison benchmark of the effectiveness of each of the four neutralizations. Participants in the pilot study were undergraduate psychology students in the Netherlands, whereas participants in Studies 1 and 2 were Amazon Mechanical Turk (MTurk) workers. MTurk workers were only selected if they (a) were located in North America (the United States or Canada) and (b) if their Human Intelligence Task approval rate > 95%. The scenarios of the pilot study and Study 1 were similar with minor differences, in part because the original Dutch-language vignette had to be translated into English. Their similarity allowed us to examine the generalizability of the outcomes across two rather dissimilar populations: mostly young Dutch students compared to the demographically more diverse MTurk workers. Study 2 utilized a different scenario. Here, instead of describing a relatively active type of misbehavior requiring personal initiative—as was the case in the pilot study and Study 1—we described a relatively passive "nonact," that is, the omission of normative behavior. This allowed us to test the robustness of any observed differences across scenarios.

Pilot Study

Method

Data Collection and Sample

We assigned 230 undergraduate psychology students (M_{age} = 20.73, SD_{age} = 3.31; 151 females, 64 males, 15 undisclosed) at a Dutch university to one of eight conditions, one dishonest/no neutralization condition, four different neutralization conditions, two other conditions included for exploratory purposes (see Footnote 2), and an honest control condition. All scenarios can be found in Appendix A.

Measures

We developed a scenario describing a male student who finds a wallet lying on the pavement, containing 1,000 euros. Before handing the wallet over to the police, the protagonist takes out one 100-euro banknote to keep for himself. Five conditions differed with regard to the statement the protagonist makes when the police discover the theft. In the honest control condition, the protagonist did not conduct theft (honest condition). Appendix A includes the six variants of the scenario as provided to the respondents. For the sake of clarity, Table 2 gives a brief overview of the six conditions and the size of the subsamples, including those for Studies 1 and 2.

To determine the level of attributed morality and sociability, we used items developed by Leach et al. (2007). More specifically, to assess perceived morality, we asked respondents to indicate to what degree they believed the protagonist of the scenario to be *honest*, *sincere*, and *trustworthy* (α = .89). To assess perceived sociability (cf. Eckes, 2002; Heflick et al., 2011; Russell & Fiske, 2008), we asked them to characterize him using the items *kind*, *warm*, and *friendly* (α = .91). Responses could be given on a 7-point Likert scale.[3]

To assess the evaluation of behavior, we asked respondents what they thought of the protagonist's behavior in terms of six semantic differential items, using a 7-point scale (negative–positive; serious–not serious; not social–social; insincere–sincere; dishonest–honest; harmful–harmless; α = .91).[4]

To measure the preferred punishment for the protagonist, we asked respondents, "What penalty do you think would be appropriate for the student?" and presented them with a 7-point scale anchored by "no penalty" and "the severest penalty possible."

To check whether our manipulation was successful, we asked respondents to answer four questions about the neutralization the protagonist used. Respondents were asked to indicate, on 7-point scale ranging from (*not at all*) to (*very much*) to what extent they thought that what the protagonist said after his confession showed that he (a) feels sorry for his behavior (apology), (b) brings up an excuse for his behavior (excuse), (c) thinks the goal justifies the means (consequentialist justification), and (d) thinks he had a right to do what he did (deontological justification).

---

[2] All studies included a control condition in which the protagonist behaves honestly and hands the wallet with its entire contents over to the police, which allowed us to assess the perception the participants had of the act of withholding money from the owner. This condition had no relevance for the purpose of comparing the impact of various neutralizations. In the pilot study, we subjected the participants to two additional experimental conditions, for exploratory purposes. These contained an account that included what are sometimes thought to be neutralizations but which fall outside our definition: a trivialization and a dissociation. These two strategies do not involve taking a clear stand vis-à-vis the norm that has been broken but instead involve an attempt to muddy the picture. First, trivializations frame the act in terms that make it seem more innocuous, by concealing the extent of a discrepancy between the norm and the behavior that breaches that norm (Bandura, 1990; Joule & Martinie, 2008; Martinie & Fointiat, 2006; Simon et al., 1995). It has also been referred to as "denial of injury" (Sykes & Matza, 1957), justification by comparison (Cromwell & Thurman, 2003), "minimizing, ignoring or misconstruing the consequences," "euphemistic labeling," and "palliative comparison" (Bandura, 1990). It can also take the form of humor (e.g., Lindsay, 2014) or the claim that "everybody is doing it" (Coleman, 1987). In our scenario, we had the protagonist state that he thinks he should not have done it but also that he can imagine more serious crimes being committed. Second, dissociation or disidentification (Aronson et al., 1995) has been found to be a means to reduce cognitive dissonance without having to change the discrepant attitude or behavior, simply by reaffirming positive self-conceptions. We operationalized it by having the protagonist in the scenario say that he should not have done it but that his behavior does not represent him as a person, adding, "I was not being myself that night." This type of account can be considered a form of excuse-making. It does not contest the validity of the norm that was breached but declares that the "real me," being different from the temporary "not really me" who breached the norm, does not carry responsibility for it. These conditions were no longer included in Studies 1 and 2.

[3] In line with the original instrument by Fiske et al. (2002) and for the sake of comprehensiveness, we also measured perceived competence with the items suggested by Leach et al. (2007). We observed no significant differences between the conditions on this measure. Since we did not develop any expectations with regard to perceived competence, we do not present these findings here, but they are available upon request. Additionally, all studies also included a brief measure of power perceptions, which was included for exploratory purposes. We found no effects of any of the manipulations on that measure.

[4] For the pilot study, which relied on a Dutch-speaking sample, the scenario was presented in Dutch with the protagonist being a nameless male student.

Table 2
Overview of the Six Conditions

Condition	Neutralization used	*N* (pilot)	*N* (study 1)	*N* (study 2)
1	Honest (control)	39	72	67
2	No neutralization (control)	38	62	65
3	Apologizing	38	57	70
4	Excuse-making	38	57	62
5	Consequentialist justification	38	55	65
6	Deontological justification	39	51	69
Total		230	354	398

Results

Manipulation Checks

An overview of means per condition can be found in Appendix Table D1. One-way analyses of variance (ANOVAs) showed that the neutralizations successfully influenced whether participants' thought that (a) the protagonist felt sorry about his behavior, $F(3, 149) = 48.73$, $p < .001$, $\eta^2 = .50$, (b) used an excuse, $F(3, 149) = 11.71$, $p < .001$, $\eta^2 = .19$, (c) thought the goal justified the means, $F(3, 149) = 9.81$, $p < .001$, $\eta^2 = .16$, and (d) thought he had the right to do what he did, $F(3, 149) = 16.36$, $p < .001$, $\eta^2 = .25$. Post hoc tests with Holm correction showed that participants in the apology condition perceived the protagonist to feel sorrier about his behavior than participants in the two justification conditions (both $p < .001$). However, participants in the apology and excuse conditions did not differ in the way they perceived the protagonist to feel sorry ($p = .27$). Participants in the excuse condition perceived the protagonist to have used an excuse to a greater extent than those in the apology condition ($p < .001$). However, there was no significant difference between the excuse condition and the two justification conditions. The neutralization conditions also differed in the degree to which participants perceived the protagonist to think that the goal justifies the means. Post hoc tests showed that in both the consequentialist and deontological justification conditions, participants perceived the protagonist to think that the goal justified the means significantly more than in the apology condition (both $p < .001$). All other conditions did not differ significantly from each other. For the last manipulation check, we tested whether the neutralizations differed significantly in the way that the participants perceived the protagonist to think that he had the right to do what he did. The post hoc *t* test showed that only the apology condition differed significantly from all other conditions (all $p < .001$), while the excuse, consequentialist justification, and deontological justification conditions did not differ significantly from each other. To sum up, the manipulation checks showed that the neutralization conditions differed significantly along the four components of the manipulation check, although not all follow-up contrasts were significant.

Main Analysis

Table 3 displays descriptive statistics and bivariate correlations. Means and standard deviations for each condition are depicted in Table 4. We first conducted one-way ANOVAs for every dependent variable in our study. The ANOVAs for morality and sociability were significant (morality: $F[1.21, 224] = 28.21$, $p < .001$, $\eta^2 = .37$; sociability: $F[1.13, 224] = 14.17$, $p < .001$, $\eta^2 = .24$), whereas the ANOVAs for evaluation of behavior and punishment were not (evaluation of behavior: $F[1.29, 186] = 1.28$, $p = .28$, $\eta^2 = .02$; punishment: $F[1.34, 186] = 1.67$, $p = .16$, $\eta^2 = .03$). The results of these tests indicate that participants' perception of morality and sociability depended on the account the protagonist gave, whereas the perception of punishment and the evaluation of behavior did not. This means that we did not find evidence for Hypotheses 1c and 1d, and no further contrast analyses will be conducted to test these hypotheses. For the significant ANOVAs, we subsequently proceeded to test our hypotheses by examining the specified contrasts summarized in Table 5.

Our hypotheses predicted that an individual who admitted wrongdoing via an apology or excuse would be perceived as more moral (H1a) and more sociable (H1b), which our results support, for morality, $t(224) = 3.75$, $p < .001$, $d = 0.56$; for sociability, $t(224) = 4.51$, $p < .001$, $d = 0.69$. However, our predictions that admissions of wrongdoing would also affect condemnation (H1c) and punishment (H1d) were not supported by the data, given that the ANOVAs were not significant. Thus, although there were no significant effects on the evaluation of behavior or the severity of the advocated punishment, protagonists who admitted having done wrong (by using apologies or excuses) were seen as more moral and sociable than those who did not admit having done wrong (by using consequentialist or deontological justifications).

Next, we explored whether it made a difference if the neutralization used was an apology or an excuse (Contrast 2; responsibility acknowledged vs. denied). This was not the case. Offering an apology or an excuse had no effect on perceived morality, $t(224) = 0.57$, $p = .57$, $d = -0.12$, or perceived sociability, $t(224) = 1.39$, $p = .16$, $d = -0.33$.

Finally, we examined potential differences between the two types of justifications (Contrast 3; consequence-dependent vs. consequence-independent). We reasoned that a deontological justification might lead to more negative evaluations than a consequentialist one. The protagonist who used a consequentialist justification was indeed perceived as more moral, $t(224) = -2.09$, $p = .04$, $d = 0.44$, but not as more sociable, $t(224) = -1.71$, $p = .08$, $d = 0.35$ than the one who used a deontological justification.

For exploratory purposes, we also investigated whether the form of neutralization used influenced the perceptions of the protagonists'

Table 3
Descriptive Statistics for Study Variables Across Conditions (Pilot)

Variable	*M*	*SD*	1	2	3
1. Morality	4.55	1.53	—		
2. Sociability	4.89	1.28	.77*	—	
3. Evaluation of behavior	3.85	1.19	.66*	.57*	—
4. Punishment	2.28	1.35	−.04	−.06	−.01

Note. $N = 186$. Higher scores for morality and warmth indicate a higher perception of morality and sociability. Higher scores for evaluation of behavior indicate a more positive opinion regarding the behavior of the protagonist. Higher scores for punishment indicate a preference for a harsher penalty.
* $p < .01$.

Table 4
Means and Standard Deviations (in Brackets) of Variables for Each Condition (Pilot)

Condition	Variable			
	Morality	Sociability	Evaluation of behavior	Punishment
Honest (control)	6.49 (.64)	5.99 (.77)	—	—
No neutralization (control)	3.88 (1.26)	4.74 (1.19)	3.65 (1.05)	2.24 (1.24)
Apologizing	4.52 (1.26)	4.88 (0.98)	3.98 (1.01)	1.92 (1.12)
Excuse-making	4.68 (1.36)	5.24 (1.21)	4.04 (1.46)	2.21 (1.42)
Consequentialist justification	4.15 (1.30)	4.46 (1.26)	3.99 (1.18)	2.34 (1.26)
Deontological justification	3.57 (1.31)	4.02 (1.25)	3.58 (1.20)	2.69 (1.59)

power and the participants' judgment of the neutralization. Results can be found in Appendix C.

Discussion and Introduction to Study 1

The pilot study provided initial evidence that different neutralizations used after a transgression can result in different social perceptions of observers. The objective of Study 1 was to replicate the hypothesized findings of the pilot study pertaining to the differential effects of neutralizations that do versus do not acknowledge wrongdoing (Contrast 1), test their robustness and generalizability in a different sample, and provide a confirmatory test of the observation in the pilot that perpetrators who use consequentialist justifications are perceived more favorably than those who use deontological justifications (Contrast 3). Thus, based on the pilot results, we hypothesized that respondents have a more positive view of a protagonist who uses a consequentialist justification than of one who uses a deontological justification, with regard to their perception of the individual's morality (Hypothesis 2a) and sociability (Hypothesis 2b). Moreover, even though we did not find significant overall effects of condition on evaluation of behavior and preferred severity of punishment, we consider the possibility that a better powered study might still provide support for our hypotheses that evaluation of his behavior (Hypothesis 2c) and preferred severity of punishment (Hypothesis 2d) might also be impacted by justification type.

Method

Data Collection and Sample

Using G*Power 3.1 software for a power analysis based on the mean effect size of Cohen's *d* obtained in the pilot study (.505) and a statistical power of .80 ($1 - \beta$) revealed a recommended minimum sample size of 126. We collected data from 504 participants through Amazon's Mechanical Turk, deliberately oversampling to account for possible incomplete answers or incorrect answers to control questions. (We included two control questions. One asked, "How much money was in the wallet?" The other was an instruction check asking participants to indicate if they were paying attention.) Again, participants were assigned to one out of seven conditions, one no neutralization condition, four different neutralization conditions, one other neutralization condition included for exploratory purposes (see Footnote 2), and a control condition. We removed 88 respondents for failing one or both control questions, and we only included four of the five neutralization conditions, resulting in a final sample of 355 participants.

Due to an oversight on our part, the study did not register gender, age, and nationality. However, based on several studies revealing the general profile of Amazon Mechanical Turk workers (Difallah et al., 2018; Huff & Tingley, 2015; Levay et al., 2016; Ross et al., 2009), we can expect our sample to consist of slightly more women than men with an average age of around 31 years, who are mostly U.S. nationals.

Measures

We translated the scenario and survey items of the pilot study from Dutch into English. In all aspects, the scenario was equal to the one used in the pilot, except that, to appeal to a more general sample, the protagonist was no longer presented as a student but simply as "Jack."

Results

Manipulation Check

As in the pilot study, we tested whether our manipulation was successful (see Appendix Table D2, for descriptives). Again, the neutralizations differed in the way that the participants thought that (a) the protagonist felt sorry about his behavior, $F(3, 216) = 50.73$, $p < .001$, $\eta^2 = .41$; (b) used an excuse, $F(3, 216) = 65.31$, $p < .001$, $\eta^2 = 48$; (c) thought the goal justified the means, $F(3, 216) = 57.77$, $p < .001$, $\eta^2 = .45$; and (d) thought that he had the right to do what he did, $F(3, 216) = 110.4$, $p < .001$, $\eta^2 = .61$. Participants indeed perceived the extent to which the protagonist felt sorry for his behavior to be greater in the apology condition than in all other conditions (all $p < .05$). The extent to which the protagonist was thought to use an excuse was greater in the excuse condition than in the apology condition ($p < .001$). However, no significant difference was found between the excuse and consequentialist ($p = .10$) and deontological justification ($p = .10$) conditions. The extent to which the protagonist thought that the goal justified the means was significantly higher for the consequentialist condition than for the apology ($p < .001$) and excuse ($p < .001$) conditions. However, the consequentialist condition did not differ significantly from the deontological condition ($p = .76$). The extent to which the protagonist thought he had the right to do what he did showed a similar pattern. The deontological justification condition differed significantly from the apology ($p < .001$) and excuse ($p < .001$) conditions but not from the consequentialist justification

Table 5
Contrasts Computed to Test Hypotheses 1a–1d (Contrast 1) and Explore Further Differences Between Conditions (Contrasts 2, 3, and 0; All Studies)

Conditions	Contrast 0: Control vs. experimental design	Contrast 1: Wrongness admitted vs. wrongness denied	Contrast 2: Responsibility acknowledged vs. responsibility denied	Contrast 3: Consequence-dependent vs. consequence-independent
Control (honest)	−2	0	0	0
Control (no neutralization)	−2	0	0	0
Apology	1	−1	−1	0
Excuse	1	−1	1	0
Consequentialist justification	1	1	0	−1
Deontological justification	1	1	0	1

condition ($p = .70$). Thus, the four components of the manipulation check again responded to the experimental manipulation, but not all expected contrasts were significant.

Main Analysis

Table 6 shows that correlations between the four dependent variables in Study 1 are similar to those found in the pilot study. Table 7 shows the means and standard deviations for each condition.

The ANOVAs again indicated that the neutralizations used by the protagonist had a significant effect on perceived morality, $F(5, 348) = 62.57$, $p < .001$, $\eta^2 = .473$, and sociability, $F(5, 348) = 39.14$, $p < .001$, $\eta^2 = .36$, but not on the evaluation of the protagonist's behavior, $F(4, 277) = 1.31$, $p = .27$, $\eta^2 = .02$, and preferred punishment, $F(4, 277) = 0.97$, $p = .42$, $\eta^2 = .01$. Therefore, we did not find evidence for Hypotheses 1c–d and 2c–d. Therefore, no further contrast analyses will be conducted to test these hypotheses. We next performed a series of contrast analyses to test Hypotheses 1a–1b and 2a–2b.

Hypotheses 1a–1b predicted that an individual who admitted wrongdoing via an apology or excuse would be perceived as more moral (H1a) and more sociable (H1b), which our results support, for morality, $t(348) = 3.94$, $p < .001$, $d = 0.49$; for sociability, $t(348) = 3.11$, $p = .002$, $d = 0.39$. However, our predictions that admissions of wrongdoing would also affect condemnation (H1c) and punishment (H1d) were not supported in the data, given that the ANOVAs were not significant. These results corroborate those of Study 1.

Next, we explored potential differences between the use of an apology and an excuse (Contrast 2), which—as in the pilot—yielded no significant effects for perceived morality, $t(348) = -0.17$, $p = .863$, $d = 0.03$; for perceived sociability, $t(348) = -0.34$, $p = .74$, $d = -0.06$. There were no differences for evaluation of behavior and preferred level of punishment.

Finally, we tested Hypotheses 2a–2b (Contrast 3). The difference observed in the pilot study between the use of the two types of justification with regard to perceived morality was not replicated, $t(348) = -1.61$, $p = .11$, $d = 0.29$, providing no support for Hypothesis 2a that offering a deontological justification would be judged less favorable than offering a consequentialist one. Justifying behavior by referring to its consequences did produce slightly stronger perceptions of sociability than justifying by appealing to rights, although this effect did not reach the threshold of $p < .05$, $t(348) = -1.85$, $p = .065$, $d = 0.32$, needed to support Hypothesis 2b. As in the pilot study, there was no notable difference regarding the evaluation of the protagonist's behavior and the preferred punishment when comparing the responses of the two conditions (Hypotheses 2c–2d).

As in the pilot, we also investigated exploratorily whether the form of neutralization used influenced the perceptions of the protagonists' power and the judgment of the protagonists' neutralization used. Results can be found in Appendix C.

Table 6
Descriptive Statistics for Study Variables Across All Conditions (Study 1)

Variable	*M*	*SD*	1	2	3
1. Morality	4.22	1.39	—		
2. Sociability	4.64	1.73	.84**	—	
3. Evaluation of behavior	3.12	1.23	.60**	.46**	—
4. Punishment	2.22	1.22	−.41**	−.33**	−.34**

Note. $N = 283$.
** $p < .01$.

Discussion and Introduction to Study 2

Study 1 lends additional credence to the idea that different neutralizations employed by perpetrators after a transgression have differential effects on observers' social judgments. Although it did not fully replicate the results of the pilot study, Study 1 does provide further support for a difference in reactions between witnessing a protagonist admitting wrongdoing and one not doing so in terms of perceived morality and sociability. In addition, when contrasting reactions to the two justification conditions, small differences in the expected direction and consistent with the pattern of the pilot study were observed, although these effects were not statistically significant.

The aim of Study 2 was to establish the generalizability of the observed patterns by testing them in a different context. In the scenario used in the pilot and Study 1, the deviant behavior required active initiative in that the 100-dollar banknote had to be taken out of the wallet. This begs the question of whether similar or different effects might be observed when a transgression consists of an omission instead of an act. It has often been observed that omissions leading to damage are considered less egregious than purposeful acts having the same effect (e.g., Cushman et al., 2006). Presumably, this is because it is less certain that wrongful intent is present within a passive individual than within an active one.

Table 7
Means and Standard Deviations (in Brackets) of Variables for Each Condition (Study 1)

Condition	Variable			
	Morality	Sociability	Evaluation of behavior	Punishment
Honest (control)	6.49 (0.86)	6.23 (0.88)	—	—
No neutralization (control)	3.40 (1.37)	4.19 (1.12)	2.95 (1.08)	2.18 (1.09)
Apologizing	4.05 (1.36)	4.43 (1.22)	3.14 (1.31)	2.05 (1.11)
Excuse-making	4.01 (1.25)	4.50 (1.03)	3.27 (1.2)	2.40 (1.51)
Consequentialist justification	3.55 (1.48)	4.19 (1.29)	3.34 (1.32)	2.09 (1.19)
Deontological justification	3.16 (1.27)	3.79 (1.21)	2.92 (1.13)	2.39 (1.33)

Against this backdrop, we set out to examine whether the general patterns observed in the pilot, and Study 1 would generalize to transgressions involving inaction rather than action.

Method

Data Collection and Sample

As in Study 1, we recruited participants through Amazon's Mechanical Turk. Using G*Power 3.1 software for a power analysis based on the mean effect sizes of Cohen's *d* obtained in Studies 1–2 (.371) and a statistical power of .80 (1—β) revealed a recommended minimum sample size of 232. We deliberately oversampled, and after removing 19 participants who had failed to answer correctly one or both of the same control questions as in Study 2 or who had filled out the questionnaire incompletely, a sample of 398 participants remained for analysis. Again, due to an oversight, demographic data were not logged. The same general demographic observations apply here as made in Study 1.

Measures

We developed a new scenario, which can be found in Appendix B. The protagonist, again a man named "Jack," fails to return a 10-dollar banknote, which he was erroneously given as change when buying a cup of coffee and a muffin at Starbucks. He either goes back the next day and hands back the 10 dollars while offering an apology, or he meets with a friend to tell him the story and why he had not returned the 10 dollars, using either an excuse, a consequentialist justification, or a deontological justification. The survey items used to measure perceived morality and sociability, evaluation of behavior, and preferred punishment were the same as in the previous studies.

Results

Manipulation Check

Similar to the pilot study and Study 1, we first tested whether our manipulation was successful (see Appendix Table D3). Four one-way ANOVAs showed that the neutralizations (apology, excuse, consequentialist justification, deontological justification) differed in the extent to which the protagonist (a) felt sorry for his behavior, $F(3, 262) = 74.13$, $p < .001$, $\eta^2 = .46$; (b) used an excuse, $F(3, 262) = 52.33$, $p < .001$, $\eta^2 = .37$; (c) thought the goal justified the means, $F(3, 262) = 47.01$, $p < .001$, $\eta^2 = .35$; and (d) thought he had the right to do what he did, $F(3, 262) = 65.17$, $p < .001$, $\eta^2 = .43$. To gain more information about how the neutralizations differed, we conducted post hoc pairwise *t* test with Holm correction. The tests showed that indeed, in the apology condition, the protagonist was perceived to feel sorry for his behavior to a greater extent than in the other three conditions (all $p < .001$). The *t* tests also showed that the participants from the excuse condition perceived the protagonist to be using excuses to a greater extent than in the apology condition ($p < .001$). However, the excuse condition neither differed from the consequentialist justification condition ($p = 0.85$) nor the deontological justification condition ($p = 0.85$) in the extent to how much the protagonist was seen as having used an excuse. When the protagonist used a consequentialist justification, participants perceived the protagonist to think that the goal justified the means to a greater extent than when he used an excuse ($p < .001$) or an apology ($p < .001$). There was no significant difference between using a consequentialist justification and using a deontological justification ($p = .12$). Similarly, when the protagonist used a deontological justification, the protagonist was perceived as thinking that he had the right to do what he did to a greater extent than if he used an apology ($p < .001$) or excuse ($p < .001$). Again, there was no significant difference between the perception when using a consequentialist justification versus using a deontological justification ($p = .43$). In short, as in the previous studies, the manipulations affected the four dimensions of the manipulation check, but not all expected contrasts were significant.

Main Analysis

Table 8 displays correlations between the four dependent variables studied in Study 2. Table 9 shows the means and standard deviations for each of the six conditions.

One-way ANOVAs indicated participants' judgments of the protagonists morality, $F(5, 392) = 91$, $p < .001$, $\eta^2 = .54$; sociability, $F(5, 392) = 65.06$, $p < .001$, $\eta^2 = .45$; and evaluation of behavior, $F(4, 326) = 2.80$, $p = .03$, $\eta^2 = .03$, were dependent on the neutralization the protagonist used, while the participant's preferred punishment for the protagonist was not, $F(4, 326) = 2.21$, $p = 07$, $\eta^2 = .03$. The nonsignificant ANOVA indicates that we did not find evidence for Hypotheses 1d and 2d. (Thus, no further contrast analyses will be conducted to test these hypotheses.) Next, as in the previous studies, we performed a series of planned comparisons for the significant ANOVAs to test our hypotheses.

Hypotheses 1a–1c predicted that an individual who admitted wrongdoing via an apology or excuse would be perceived as more moral (H1a) and more sociable (H1b) and that the behavior would consequentially be evaluated as more positive (H1c).

Table 8
Descriptive Statistics for Study Variables Across All Conditions (Study 2)

Variable	*M*	*SD*	1	2	3
1. Morality	3.58	1.90	—		
2. Sociability	3.98	1.62	.87**	—	
3. Evaluation of behavior	2.91	1.23	.72**	.57**	—
4. Punishment	2.83	1.54	−.26**	−.28**	−.24**

Note. $N = 331$.
** $p < .01$.

These hypotheses were supported by the results, for morality, $t(392) = 7.07$, $p < .001$, $d = 0.82$; for sociability, $t(392) = 7.70$, $p < .001$, $d = 0.91$; for evaluation of behavior, $t(392) = 2.17$, $p = .031$, $d = 0.26$. However, our prediction that admissions of wrongdoing would also affect punishment (H1d) were not supported in the data, given that the ANOVA was not significant.

Next, we explored potential differences in outcomes depending on whether the neutralization used was an apology or an excuse (Contrast 2). As in the pilot and Study 1, this made no significant difference with regard to perceived morality, $t(392) = -1.71$, $p = .09$, $d = 0.28$, perceived sociability, $t(392) = -1.75$, $p = .081$, $d = 0.31$, or evaluation of behavior, $t(326) = 0.04$, $p = .96$, $d = -0.01$.

Finally, we tested Hypotheses 2a–2c by contrasting the effects of using a consequentialist justification with the effects of using a deontological justification (Contrast 3). We found support for Hypothesis 2b that the protagonist who used consequential justification was perceived as more sociable than the one who used deontological justification, $t(392) = -2.04$, $p = .04$, $d = 0.32$. For the remaining hypotheses, we did not find support, as the differences with regard to perceived morality, $t(392) = -1.72$, $p = .09$, $d = 0.28$, or evaluation of behavior, $t(326) = -1.76$, $p = .08$, $d = 0.28$, did not reach significance, although the effect evaluation of behavior was in the expected direction.

Analogue to the pilot study and Study 1, we again explored whether the neutralization used influenced the perception of the protagonists' power and the judgment of the neutralization used. Results are presented in Appendix C.

Internal Meta-Analysis

Whereas some hypotheses were consistently supported across all three studies, support for other hypotheses was more varied. We therefore conducted an internal meta-analysis to obtain a more reliable estimate of the effects of interest. Specifically, we meta-analytically tested differences between using an apology or an excuse on the one hand and not admitting wrongdoing by using a justification (Contrast 1; Hypotheses 1a–1d), using an apology or an excuse (Contrast 2), and using a consequentialist versus a deontological justification (Contrast 3; Hypotheses 2a–2d).

To conduct this meta-analysis, we used the statistical software R. We calculated Cohen's *d* and the square root of its sampling variance, which is Cohen's *d*'s standard error. In light of the small sample of studies, we used fixed effect models to calculate the overall effect sizes. Tables 10–12 present Cohen's *d*'s of the four dependent variables and their standard errors across the three studies for the three contrasts. The tables also include heterogeneity statistics. A high likelihood of heterogeneity was observed in only two instances; both pertain to the pooled effect on sociability evaluations (Contrasts 1 and 2). It is possible that sociability evaluations are—more so than morality and behavior evaluations—contingent on other factors. Forest plots of these Cohen's *d*'s are depicted in Figures 1–3. Each figure shows a different contrast regarding perceived level of morality and sociability, evaluation of behavior, and severity of punishment for the three studies, with the diamond showing the overall Cohen's *d* (and 95% confidence interval) yielded by an internal meta-analysis across the three studies.

The results of the meta-analysis comparing the use of neutralizations that do versus do not acknowledge wrongdoing (Contrast 1) provide support for Hypotheses 1a–1d. The aggregated Cohen's *d*'s for morality, sociability, evaluation of behavior, and preferred severity of punishment were all significant. Overall, compared to using a consequentialist or deontological justification, using an apology or an excuse resulted in higher levels of perceived morality (0.64 [CI: 0.48, 0.80], $p < .001$) and sociability (0.67 [CI: 0.51, 0.83], $p < .001$), a more positive evaluation of behavior (0.17 [CI: 0.02, 0.33], $p = .03$), and a lower degree of preferred punishment (−0.22 [CI: −0.38, −0.07], $p < .01$; see Figure 1).

The meta-analysis of the effects of using an apology versus an excuse (Contrast 2) did not yield significant differences on three of the four dependent variables. The aggregated Cohen's *d*'s for morality, sociability, evaluation of behavior, and preferred severity of punishment were not significant (0.09 [95% CI: −0.12, 0.31], $p = .40$; 0.03 [95% CI: −0.19, 0.25], $p = .92$; −0.05 [95% CI: −0.27, 0.17], $p = .65$; and −0.15 [95% CI: −0.37, 0.07], $p = .19$, respectively).

The meta-analysis of the effects of using consequentialist versus deontological justifications (Contrast 3) revealed a significant aggregated Cohen's *d* across the three studies for the effects on perceived morality (0.32 [CI: 0.10, 0.54], $p < .01$), perceived sociability (0.33 [CI: 0.11, 0.54], $p < .01$), and evaluation of

Table 9
Means and Standard Deviations (in Brackets) of Variables for Each Condition (Study 2)

Condition	Variable			
	Morality	Sociability	Evaluation of behavior	Punishment
Honest (control)	6.42 (0.93)	6.08 (0.92)	—	—
No neutralization (control)	2.46 (1.28)	3.06 (1.21)	2.67 (1.00)	2.97 (1.48)
Apologizing	3.9 (1.27)	4.44 (1.11)	3.13 (1.16)	2.53 (1.38)
Excuse-making	3.52 (1.54)	4.07 (1.30)	3.14 (1.28)	2.53 (1.47)
Consequentialist justification	2.77 (1.45)	3.33 (1.40)	3 (1.37)	3 (1.69)
Deontological justification	2.39 (1.28)	2.9 (1.25)	2.63 (1.37)	3.12 (1.61)

Table 10
Cohen's ds and Associated Standard Errors for Contrast 1 (Admitting vs. Not Admitting Wrongness) in Each Study, Plus Heterogeneity Statistics (Cochrane's Q)

	Morality		Sociability		Evaluation of behavior		Preferred punishment	
Dependent variable	Cohen's *d*	SE_d	Cohen's *d*	SE_d	Cohen's *d*	SE_d	Cohen's *d*	SE_d
Pilot	.56	.16	.69	.17	.19	.16	−.33	.16
Study 1	.49	.14	.39	.14	.05	.13	−.01	.13
Study 2	.82	.13	.91	.13	.26	.12	−.34	.12
	Q	*p*	*Q*	*p*	*Q*	*p*	*Q*	*p*
Heterog.	3.32	.190	7.42	.024	1.43	.489	4.08	.129

Note. *SE* = standard error; heterog. = heterogeneity.

behavior (0.32 [CI: 0.10, 0.54], $p < .01$), supporting Hypotheses 2a–2c. We obtained no meta-analytic support for Hypothesis 2d regarding the preferred severity of punishment (−0.17 [CI: −0.39, 0.05], $p = .13$).

General Discussion

The present studies were conducted to assess whether the type of neutralization used by a misbehaving individual affects observers' judgments of the transgressor and the misbehavior. We distinguished between four different types of neutralizations based on three underlying theoretical dimensions: (a) whether the actor admits that the act was wrong, (b) whether the actor acknowledges full causal responsibility for the act, and (c) whether the neutralization depends for its purported validity on the consequences of the act. These three dimensions delineate four types of neutralizations (see Table 1) that have been documented previously (please refer to Footnote 1; also see Cottee, 2010; Schönbach, 2010; Scott & Lyman, 1968; Sykes & Matza, 1957): apologizing, excuse-making, consequentialist justifications, and deontological justifications.

Our main hypothesis—that observers have a more positive view of actors who admit wrongdoing than of those who do not—was supported. In all three studies, we found evidence to support Hypotheses 1a and 1b: Respondents perceived a protagonist who used either an apology or an excuse as having a significantly higher level of morality (consistent with Hypothesis 1a) and sociability (consistent with Hypothesis 1b) than one who defended his behavior by using a justification. For Hypotheses 1c and 1d regarding the evaluation of behavior and preferred punishment, respectively, evidence was mixed across studies, but an internal meta-analysis provided support: Across studies, respondents evaluated the behavior less harshly (consistent with Hypothesis 1c) and recommended a lower punishment (consistent with Hypothesis 1d) when the protagonist responded with an apology or used an excuse than when he justified his behavior.

We also explored the potential difference between acknowledging responsibility with an apology and not doing so with an excuse. In none of the studies did we find evidence that an apology or an excuse has a different effect in terms of perceived morality, perceived sociability, evaluation of behavior, or preferred punishment.

We obtained mixed evidence across the studies for our hypothesis that a consequentialist justification elicits a more positive judgment of the protagonist and his behavior than a deontological justification. The pilot study revealed that the use of a consequentialist justification resulted in a more positive perception of morality and sociability. No such evidence was obtained in Study 1. Study 2 also showed this effect for perceived morality, but not for the other outcome variables (although the differences between conditions were in the hypothesized direction). Nevertheless, outcomes of the meta-analysis yield support for the notion that the two types of justification lead to different evaluations, with the exception of preferred punishment (consistent with Hypotheses 2a–2c but not 2d).

In sum, results of three studies provide support for the assertion that the type of neutralization used upon a transgression influences observers' judgments of the transgressor and the behavior, with the most robust support pertaining to the differential effects on observers' judgments of acknowledging wrongdoing (by using an

Table 11
Cohen's ds and Associated Standard Errors for Contrast 2 (Apology vs. Excuse) in Each Study, Plus Heterogeneity Statistics (Cochrane's Q)

	Morality		Sociability		Evaluation of behavior		Preferred punishment	
Dependent variable	Cohen's *d*	SE_d	Cohen's *d*	SE_d	Cohen's *d*	SE_d	Cohen's *d*	SE_d
Pilot	−.12	.23	−.33	.23	−.05	.23	−.23	.23
Study 1	.03	.19	−.06	.19	−.10	.19	−.26	.19
Study 2	.28	.18	.31	.18	−.01	.17	.00	.17
	Q	*p*	*Q*	*p*	*Q*	*p*	*Q*	*p*
Heterog.	2.05	.359	5.09	.078	0.12	.939	1.23	.540

Note. *SE* = standard error; heterog. = heterogeneity.

Table 12
Cohen's ds and Associated Standard Errors for Contrast 3 (Consequentialist vs. Deontological Justification) in Each Study, Plus Heterogeneity Statistics (Cochrane's Q)

	Morality		Sociability		Evaluation of behavior		Preferred punishment	
Dependent variable	Cohen's *d*	SE_d	Cohen's *d*	SE_d	Cohen's *d*	SE_d	Cohen's *d*	SE_d
Pilot	.44	.23	.35	.23	.34	.23	−.24	.23
Study 1	.29	.20	.32	.20	.34	.20	−.24	.20
Study 2	.28	.17	.32	.17	.28	.17	−.07	.17
	Q	*p*	*Q*	*p*	*Q*	*p*	*Q*	*p*
Heterog.	0.35	.839	0.10	.993	0.07	.965	0.56	.756

Note. *SE* = standard error; heterog. = heterogeneity.

apology or an excuse) versus not acknowledging wrongdoing (by using a consequentialist or deontological justification). Meta-analytic evidence further indicates that consequentialist justifications elicit more favorable judgments than deontological judgments, although this effect was not borne out in all individual studies.

Limitations and Suggestions for Future Research

One limitation to the generalizability of our findings might lie in the use of hypothetical scenarios. Judging other people's behavior from a situational and temporal distance, and in the knowledge that the described behavior is "just a story," is likely to differ from perceptions of behavior in real time and real life (e.g., Straughan, 1975). However, it is important to note that in real life too, people are often confronted with indirect descriptions of behavior, which they subsequently form an opinion about. For example, people may read about the misbehavior of a celebrity or politician online, which informs their attitudes about that person. Future research could examine whether the effects observed here are similar or different when people are confronted with transgressions and neutralizations at the moment.

Another limitation pertains to the lack of insight into the mechanism responsible for the effects that we observed. For example, what causes justifications to foster more negative evaluations than excuses? One possibility, which we already touched upon in the introduction, is that justifications are perceived as signaling a higher likelihood of recidivism as well as the possibility of a more general erosion of norms because no wrongness is admitted. This might be especially problematic in the case of widely shared norms. Suggestive evidence for this possibility comes from previous research on the intrapersonal correlates of different neutralizations discussed above, but clearly more direct evidence is needed. Future research could examine these and other possible underlying processes.

A related limitation is that the patterns of results observed on our manipulation checks reveal that respondents found it difficult to differentiate between some of the experimental conditions in terms of the degree to which the protagonist employed different types of neutralizations. Although the observed patterns were generally in the expected direction (see Appendix D) and all four components of the manipulation checks were significantly affected by the manipulations in all three studies, not all contrasts between conditions were significant. Nonetheless, we did observe theoretically meaningful patterns on our substantive dependent variables, suggesting that participants did pick up on the more subtle differences between the conditions. A possible explanation is that participants registered the different neutralizations of the protagonist in a more holistic or intuitive manner, so that they were able to develop different perceptions of the protagonist across conditions without being able to articulate exactly how the protagonist's behavior corresponded with the specific, rather technical dimensions underlying the manipulations. In other words, the impact of reading about neutralization use may be more intuitive than reasoned (Cushman et al., 2006), which may not have been captured well by the specific manipulation check statements. We recommend including less challenging manipulation checks in future studies to gain better insight into how people perceive the different types of neutralizations.

A final point of consideration is that the current research focuses on the effects of neutralizations following a transgression concerning theft. While the transgression as described in the scenarios used is relatively mild, theft will generally be condemned by most people. However, norm transgressions do also occur in domains in which the transgression might be viewed—at least by some—as justified (e.g., environmental activism). Future research efforts might investigate boundary conditions to the observed effects or look at the interplay between type of transgression and type of neutralization.

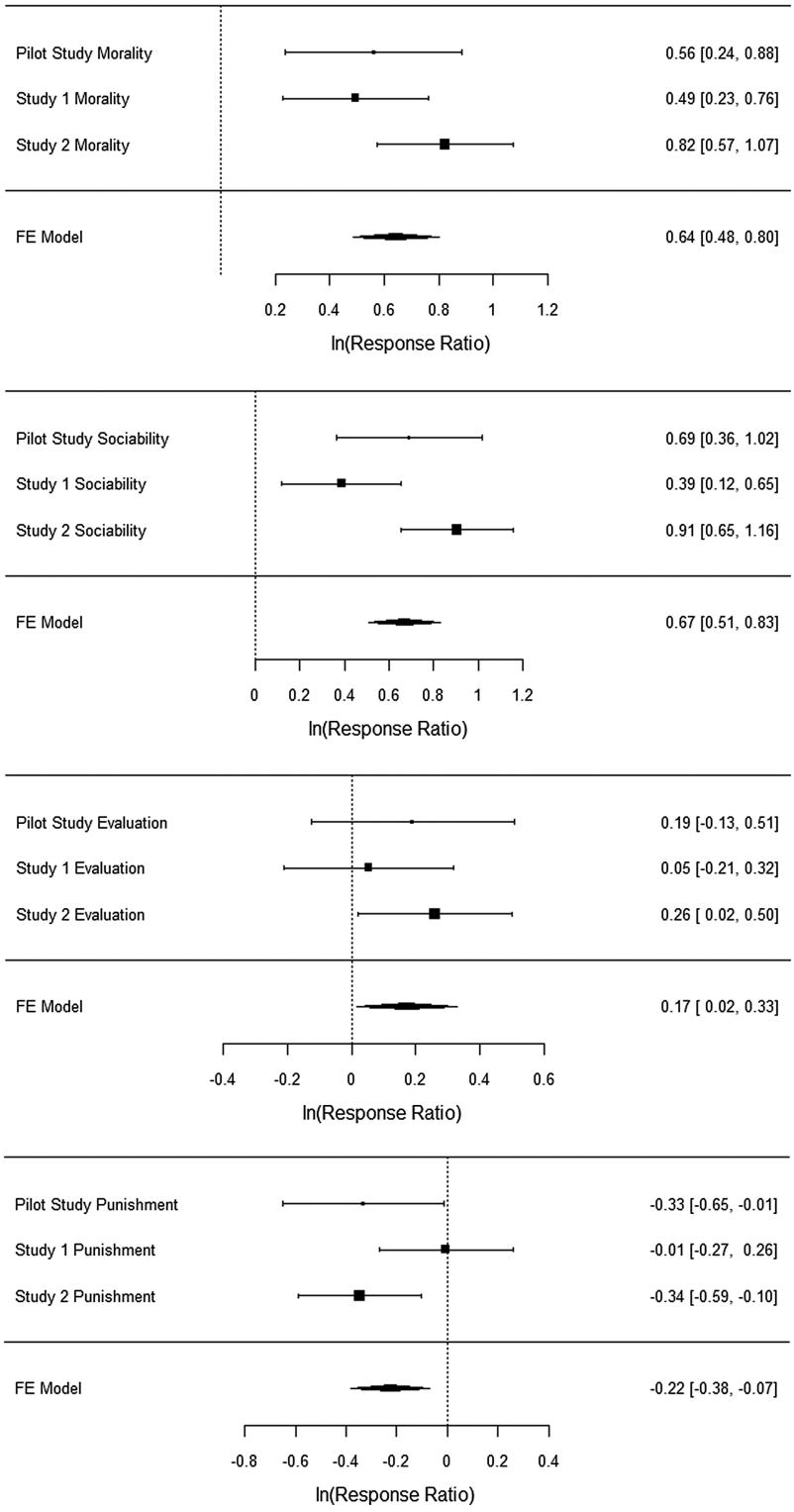
Contributions and Implications

Despite a long and rich history of scholarly interest in responses to transgressions, scientific understanding of the social consequences of using different types of neutralizations to account for transgressive behavior is still embryonic. The current work contributes to such understanding by offering a theoretically grounded approach to classifying different forms of neutralization and systematically examining the differential interpersonal effects of these neutralizations on observers' perceptions of transgressors and their behaviors.

As described earlier in this article, tests of the interpersonal effects of neutralizations are rare. Research at the intrapersonal level of analysis suggests that the type of neutralization used may predict the severity or the frequency of (future) deviant behavior (e.g., A. Barsky, 2011; A. J. Barsky et al., 2006; Copes, 2003; Cottee, 2010). The present studies complement and extend this previous work by providing insight into the *interpersonal* effects of different neutralizations. Our finding that neutralizations that involve an acknowledgment of wrongdoing (apologies and excuses) have more favorable interpersonal consequences than neutralizations that deny

Figure 1

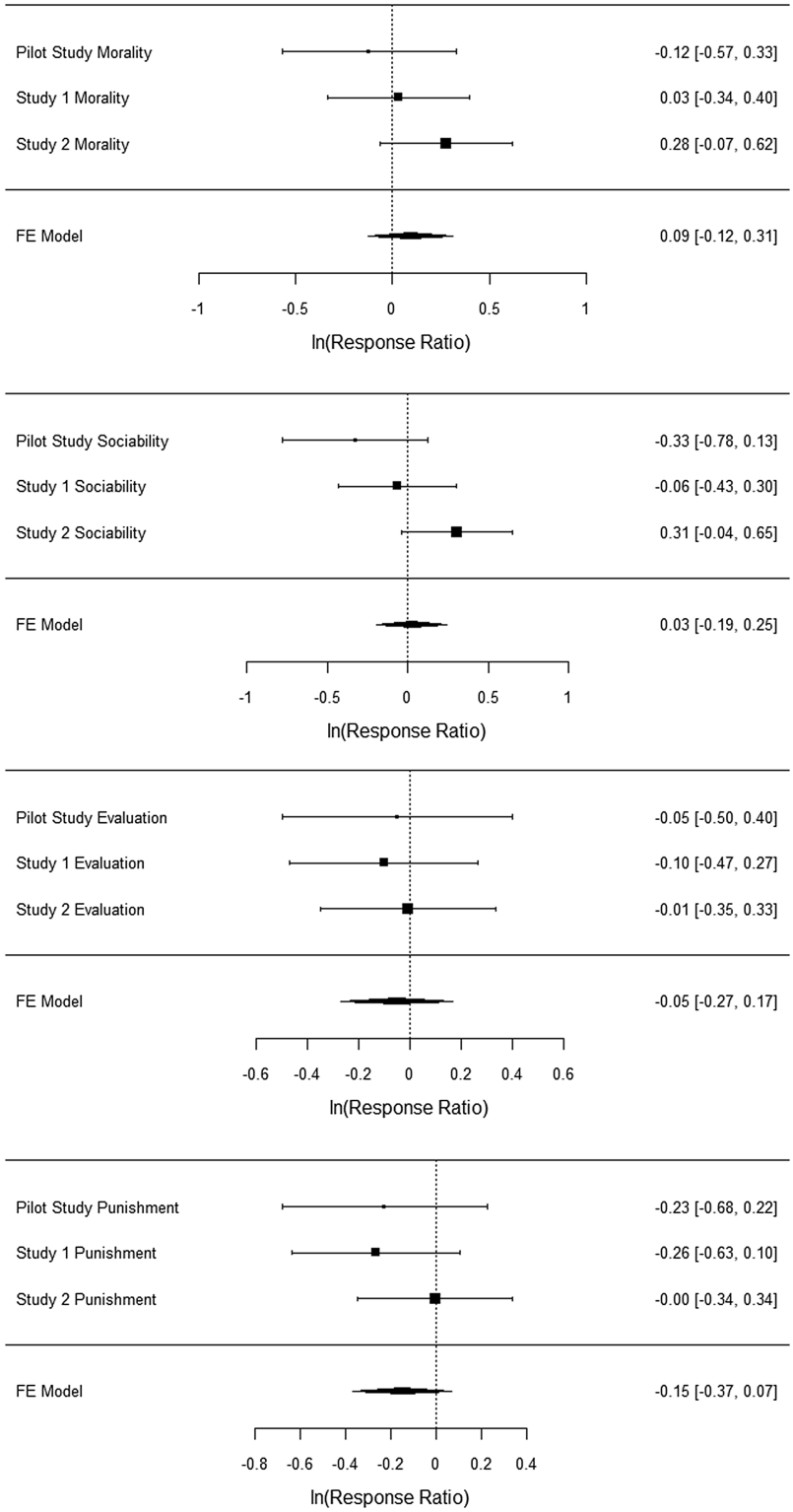
Forest Plot Displaying Cohen's ds of Contrast 1 (Admitting vs. Not Admitting Wrongness; With 95% Confidence Intervals) Across Studies



Note. FE = Fixed Effects Model.

Figure 2

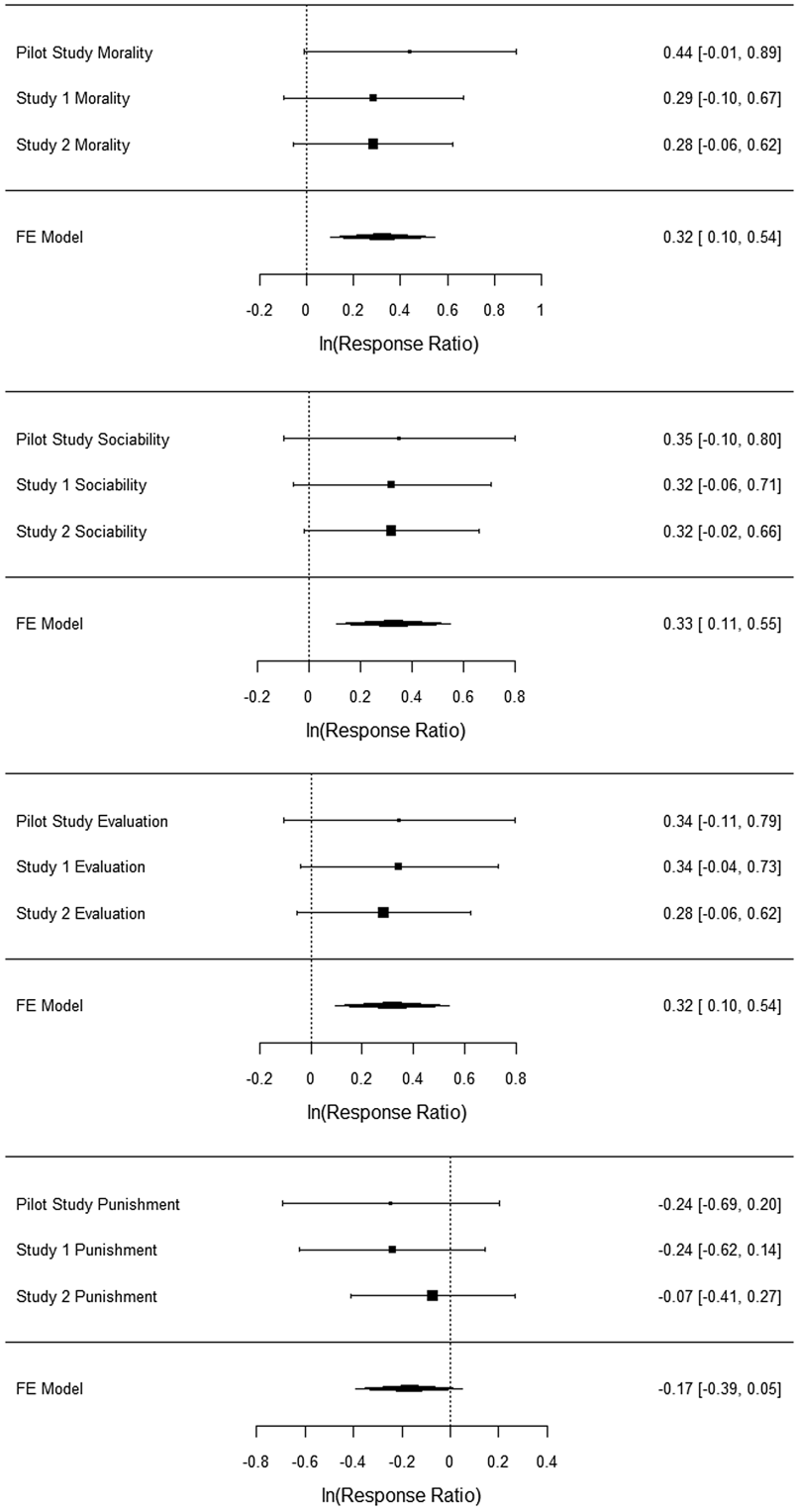
Forest Plot Displaying Cohen's ds of Contrast 2 (Apology vs. Excuse; With 95% Confidence Intervals) Across Studies



Note. FE = Fixed Effects Model.

Figure 3

Forest Plot Displaying Cohen's ds of Contrast 3 (Consequentialist vs. Deontological Justification; With 95% Confidence Intervals) Across Studies



Note. FE = Fixed Effects Model.

wrongdoing (consequentialist and deontological justifications) resonates with the emerging conclusion from work at the intrapersonal level of analysis that the former category of neutralizations is associated with less severe (recidivist) transgressions.

Conclusion

Social norms are the building blocks of durable communities. Transgressions endanger the existence and perpetuation of norms, especially when not confronted and publicly opposed. Our findings show that observers' responses to norm transgressions are shaped by the neutralizations that transgressors use to account for their behavior. In particular, the use of consequentialist or deontological justifications elicited more negative responses from observers than the use of apologies or excuses. Thus, the very neutralizations that may lower the threshold for transgressive behavior at the intrapersonal level evoke harsher social responses at the interpersonal level.

References

Agnew, R., & Peters, A. A. (1986). The techniques of neutralization: An analysis of predisposing and situational factors. *Criminal Justice and Behavior*, *13*(1), 81–97. https://doi.org/10.1177/0093854886013001005

Akers, R. L. (2009). *Social learning and social structure: A general theory of crime and deviance*. Transaction Publishers. (Original work published 1998)

Aronson, J., Blanton, H., & Cooper, J. (1995). From dissonance to disidentification: Selectivity in the self-affirmation process. *Journal of Personality and Social Psychology*, *68*(6), 986–996. https://doi.org/10.1037/0022-3514.68.6.986

Ashforth, B. E., & Anand, V. (2003). The normalization of corruption in organizations. In R. M. Kramer & B. M. Staw (Eds.), *Research in organizational behavior*, (Vol. 25, pp. 1–52). Elsevier.

Bandura, A. (1990). Selective activation and disengagement of moral control. *Journal of Social Issues*, *46*(1), 27–46. https://doi.org/10.1111/j.1540-4560.1990.tb00270.x

Barsky, A. (2011). Investigating the effects of moral disengagement and participation on unethical work behaviour. *Journal of Business Ethics*, *104*(1), 59–75. https://doi.org/10.1007/s10551-011-0889-7

Barsky, A. J., Islam, G., Zyphur, M. J., & Johnson, E. (2006). *Investigating the effects of moral Disengagement and participation on unethical work behavior* [Paper presentation]. 11th Annual Conference of the Brazilian Academy of Management, Salvador, Brazil. Insper Working Paper WPE 069/2006, pp. 1–22.

Brambilla, M., Rusconi, P., Sacchi, S., & Cherubini, P. (2011). Looking for honesty: The primary role of morality (vs. sociability and competence) in information gathering. *European Journal of Social Psychology*, *41*(2), 135–143. https://doi.org/10.1002/ejsp.744

Bryant, E., Schimke, E. B., Nyseth Brehm, H., & Uggen, C. (2017). Techniques of neutralization and identity work among accused genocide perpetrators. *Social Problems*, *65*(4), 584–602. https://doi.org/10.1093/socpro/spx026

Cechaviciute, I., & Kenny, D. T. (2007). The relationship between neutralizations and perceived delinquent labeling on criminal history in young offenders serving community orders. *Criminal Justice and Behavior*, *34*(6), 816–829. https://doi.org/10.1177/0093854806298210

Coleman, J. W. (1987). Toward an integrated theory of white-collar crime. *American Journal of Sociology*, *93*(2), 406–439. https://doi.org/10.1086/228750

Copes, H. (2003). Societal attachments, offending frequency, and techniques of neutralization. *Deviant Behavior*, *24*(2), 101–127. https://doi.org/10.1080/01639620390117200

Copley, L. (2014). What does policy have to do with it? The political economy of Latino sex trafficking in the United States. *Crime, Law and Social Change*, *62*(5), 571–584. https://doi.org/10.1007/s10611-014-9542-6

Cottee, S. (2010). Mind slaughter: The neutralizations of Jihadi Salafism. *Studies in Conflict and Terrorism*, *33*(4), 330–352. https://doi.org/10.1080/10576101003587176

Cromwell, P., & Thurman, Q. (2003). The devil made me do it: Use of neutralizations by shoplifters. *Deviant Behavior*, *24*(6), 535–550. https://doi.org/10.1080/713840271

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science*, *17*(12), 1082–1089. https://doi.org/10.1111/j.1467-9280.2006.01834.x

Difallah, D., Filatova, E., & Ipeirotis, P. (2018). *Demographics and dynamics of mechanical turk workers* [Conference session]. Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining.

Eckes, T. (2002). Paternalistic and envious gender stereotypes: Testing predictions from the stereotype content model. *Sex Roles*, *47*(3–4), 99–114. https://doi.org/10.1023/A:1021020920715

Everett, J. A., Pizarro, D. A., & Crockett, M. J. (2016). Inference of trustworthiness from intuitive moral judgments. *Journal of Experimental Psychology: General*, *145*(6), 772–787. https://doi.org/10.1037/xge0000165

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Gauthier, D. K. (2001). Professional lapses: Occupational deviance and neutralization techniques in veterinary medical practice. *Deviant Behavior*, *22*(6), 467–490. https://doi.org/10.1080/016396201753227691

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, *106*(1), 148–168. https://doi.org/10.1037/a0034726

Haines, V. J., Diekhoff, G. M., LaBeff, E. E., & Clark, R. E. (1986). College cheating: Immaturity, lack of commitment, and the neutralizing attitude. *Research in Higher Education*, *25*(4), 342–354. https://doi.org/10.1007/BF00992130

Heflick, N. A., Goldenberg, J. L., Cooper, D. P., & Puvia, E. (2011). From women to objects: Appearance focus, target gender, and perceptions of warmth, morality and competence. *Journal of Experimental Social Psychology*, *47*(3), 572–581. https://doi.org/10.1016/j.jesp.2010.12.020

Hollinger, R. C. (1991). Neutralizing in the workplace: An empirical analysis of property theft and production deviance. *Deviant Behavior*, *12*(2), 169–202. https://doi.org/10.1080/01639625.1991.9967872

Huff, C., & Tingley, D. (2015). "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research & Politics*, *2*(3), 1–12. https://doi.org/10.1177/2053168015604648

Joule, R. V., & Martinie, M. A. (2008). Forced compliance, misattribution and trivialization. *Social Behavior and Personality: An International Journal*, *36*(9), 1205–1212. https://doi.org/10.2224/sbp.2008.36.9.1205

Klenowski, P. M., Copes, H., & Mullins, C. W. (2011). Gender, identity, and accounts: How white collar offenders do gender when making sense of their crimes. *Justice Quarterly*, *28*(1), 46–69. https://doi.org/10.1080/07418825.2010.482536

Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, *93*(2), 234–249. https://doi.org/10.1037/0022-3514.93.2.234

Levay, K. E., Freese, J., & Druckman, J. N. (2016). The demographic and political composition of Mechanical Turk samples. *SAGE Open*, *6*(1). https://doi.org/10.1177/2158244016636433

Lindsay, B. C. (2014). Humor and dissonance in California's native American genocide. *American Behavioral Scientist*, *58*(1), 97–123. https://doi.org/10.1177/0002764213495034

Martinie, M. A., & Fointiat, V. (2006). Self-esteem, trivialization, and attitude change. *Swiss Journal of Psychology*, *65*(4), 221–225. https://doi.org/10.1024/1421-0185.65.4.221

Metts, S., & Cupach, W. R. (1989). Situational influence on the use of remedial strategies in embarrassing predicaments. *Communication Monographs*, *56*(2), 151–162. https://doi.org/10.1080/03637758909390256

Mills, C. W. (1940). Situated actions and vocabularies of motive. *American Sociological Review*, *5*(6), 904–913. https://doi.org/10.2307/2084524

Moore, C. (2015). Moral disengagement. *Current Opinion in Psychology*, *6*, 199–204. https://doi.org/10.1016/j.copsyc.2015.07.018

Ohbuchi, K., Kameda, M., & Agarie, N. (1989). Apology as aggression control: Its role in mediating appraisal of and response to harm. *Journal of Personality and Social Psychology*, *56*(2), 219–227. https://doi.org/10.1037/0022-3514.56.2.219

Osofsky, M. J., Bandura, A., & Zimbardo, P. G. (2005). The role of moral disengagement in the execution process. *Law and Human Behavior*, *29*(4), 371–393. https://doi.org/10.1007/s10979-005-4930-1

Ribeaud, D., & Eisner, M. (2010). Are moral disengagement, neutralization techniques, and self-serving cognitive distortions the same? Developing a unified scale of moral neutralization of aggression. *International Journal of Conflict and Violence*, *4*(2), 298–315. https://doi.org/10.4119/ijcv-2833

Robinson, S. L., & Kraatz, M. S. (1998). Constructing the reality of normative behavior: The use of neutralization strategies by organizational deviants. In R. W. Griffin, A. O'Leary-Kelly, & J. M. Collins (Eds.), *Dysfunctional behavior in organizations: Violent and deviant behavior: Monographs in organizational behavior and industrial relations*, (Vol. 23, pp. 203–220). Elsevier Science/JAI Press.

Ross, J., Zaldivar, A., Irani, L., & Tomlinson, B. (2009). *Who are the turkers? Worker demographics in amazon mechanical turk* [Tech. Rep.]. Department of Informatics, University of California.

Russell, A. M. T., & Fiske, S. T. (2008). It's all relative: Competition and status drive interpersonal perception. *European Journal of Social Psychology*, *38*(7), 1193–1201. https://doi.org/10.1002/ejsp.539

Schönbach, P. (2010). *Account episodes: The management or escalation of conflict*. Cambridge University Press. (Original work published 1990)

Scott, M. B., & Lyman, S. M. (1968). Accounts. *American Sociological Review*, *33*(1), 46–62. https://doi.org/10.2307/2092239

Simon, L., Greenberg, J., & Brehm, J. (1995). Trivialization: The forgotten mode of dissonance reduction. *Journal of Personality and Social Psychology*, *68*(2), 247–260. https://doi.org/10.1037/0022-3514.68.2.247

Straughan, R. R. (1975). Hypothetical moral situations. *Journal of Moral Education*, *4*(3), 183–189. https://doi.org/10.1080/0305724750040301

Sykes, G., & Matza, D. (1957). Techniques of neutralization: A theory of delinquency. *American Sociological Review*, *22*(6), 664–670. https://doi.org/10.2307/2089195

Van Baak, C., Hayes, B. E., Freilich, J. D., & Chermak, S. M. (2018). Honor crimes in the United States and offenders' neutralization techniques. *Deviant Behavior*, *39*(2), 187–202. https://doi.org/10.1080/01639625.2016.1266870

Van der Lee, R., Ellemers, N., Scheepers, D., & Rutjens, B. T. (2017). In or out? How the perceived morality (vs. competence) of prospective group members affects acceptance and rejection. *European Journal of Social Psychology*, *47*(6), 748–762. https://doi.org/10.1002/ejsp.2269

Van Kleef, G. A., & De Dreu, C. K. (2010). Longer-term consequences of anger expression in negotiation: Retaliation or spillover? *Journal of Experimental Social Psychology*, *46*(5), 753–760. https://doi.org/10.1016/j.jesp.2010.03.013

Van Kleef, G. A., Gelfand, M. J., & Jetten, J. (2019). The dynamic nature of social norms: New perspectives on norm development, impact, violation, and enforcement. *Journal of Experimental Social Psychology*, *84*, 1–5. https://doi.org/10.1016/j.jesp.2019.05.002

Van Kleef, G. A., Wanders, F., Stamkou, E., & Homan, A. C. (2015). The social dynamics of breaking the rules: Antecedents and consequences of norm-violating behavior. *Current Opinion in Psychology*, *6*, 25–31. https://doi.org/10.1016/j.copsyc.2015.03.013

Appendix A

Scenario Pilot and Study 1 (English Version)

1. Honest control condition—honest protagonist, no neutralization

It's late at night. Jack has been at a party and is now walking back to his car. The street is deserted. Suddenly, Jack spots an expensive-looking wallet on the ground. He picks it up and discovers that it contains ten $100 banknotes. There is nothing in the wallet that can tell Jack about the identity of the owner. Despite the late hour, Jack decides to go to the nearest police station to hand in the wallet.

A police officer shows him in, takes some notes, accepts the wallet for safekeeping, and praises Jack for his honesty. When Jack is about to leave, another officer enters the room, saying, "A man just came in to report that he had lost his wallet. If the wallet you found is his, he may want to thank you in person." Jack is then led to another room where he meets with a well-dressed middle-aged man. It is quickly established that the lost wallet and the one Jack had found are one and the same. The man is very grateful to Jack.

2. Dishonest control condition—dishonest protagonist, no neutralization

It's late at night. Jack has been at a party and is now walking back to his car. The street is deserted. Suddenly, Jack spots an expensive-looking wallet on the ground. He picks it up and discovers that it contains ten $100 banknotes. There is nothing in the wallet that can tell Jack about the identity of the owner. Despite the late hour, Jack decides to go to the nearest police station to hand in the wallet. Just before he enters the police station, Jack takes out one of the $100 banknotes and puts it in his back pocket.

A police officer shows him in, takes some notes, accepts the wallet for safekeeping, and praises Jack for his honesty. When Jack is about to leave, another officer enters the room, saying, "A man just came in to report that he had lost his wallet. If the wallet you found is his, he may want to thank you in person." Jack is then led to another room where he meets with a well-dressed middle-aged man. It is quickly established that the lost wallet and the one Jack had found are one and the same. The man is very grateful to Jack.

But then the man notices that one banknote is missing. He knows exactly how much money was in the wallet because he had shortly withdrawn $1,000 from an ATM. The police officers become

(*Appendices continue*)

suspicious and soon they discover that Jack has a banknote in his back pocket. Jack then confesses to having taken out $100 from the wallet and hands it back to the owner.

In the experimental conditions, participants saw the same text as in the dishonest control condition above, followed by one of four neutralizations:

3. Apology

Jack says, "I should not have done this. I'm really sorry. I'm offering my apologies."

4. Excuse

Jack says, "I should not have done this. But I had little control over my behavior. I wasn't thinking. It's late and I've been drinking too much."

5. Consequentialist justification

Jack says, "I did the right thing. The benefit to me, a poor guy who could really have used those $100, is greater than the harm to the owner."

6. Deontological justification

Jack says, "I did the right thing. There's such a thing as a finder's fee: the idea that the finder has a right to part of the value of something lost. I deserved to keep those $100."

Appendix B

Scenario Study 2

1. Honest control condition—honest protagonist, no neutralization

In the morning, after a sleepless night, a tired Jack waits in line for 10 min to order a coffee and muffin at Starbucks. He pays at the counter and waits for his coffee. Only then does he realize that the clerk had given him change for $20 rather than the $10 he had given her. Jack returns the $10 promptly and savors his coffee and muffin.

2. Dishonest control condition—dishonest protagonist, no neutralization

In the morning, after a sleepless night, a tired Jack waits in line for 10 min to order a coffee and muffin at Starbucks. He pays at the counter and waits for his coffee. Only then does he realize that the clerk had given him change for $20 rather than the $10 he had given her. Jack savors his coffee, muffin, and free $10.

In the experimental conditions, participants saw the same text, followed by one of four neutralizations.

3. Apology

Later that day, Jack meets with a friend and tells him what happened at Starbucks. Jack says, "I feel wrong about having kept those $10. I'm really sorry."

4. Excuse

Later that day, Jack meets with a friend and tells him what happened at Starbucks. Jack says, "I should not have kept those $10. But I had little control over my behavior. I wasn't thinking. It was very early, and I had been up all night."

5. Consequentialist justification

Later that day, Jack meets with a friend and tells him what happened at Starbucks. Jack says, "Keeping those $10 was the right thing to do. The benefit to me, a poor guy who can really use $10, is greater than the harm to Starbucks."

6. Deontological justification

Later that day, Jack meets with a friend and tells him what happened at Starbucks. Jack says, "Keeping those $10 was the right thing to do. I'm not obliged to correct the mistakes other people make. I deserved to keep the $10."

Appendix C

Exploratory Analyses on Power Perceptions and Judgment of Neutralizations

In all studies, we also investigated whether the form of neutralization used influenced participants' perceptions of the protagonists' power and judgments of the neutralization.

Pilot Study

The neutralization used did not influence the perception of the protagonists' power, $F(5, 244) = 1.35$, $p = .24$, $\eta^2 = .03$, but did influence how participants judged the excuse that was used, $F(3, 149) = 19.51$, $p < .001$, $\eta^2 = .28$. Post hoc pairwise t test showed that participants judged the apology more positively than the other neutralizations (excuse, consequentialist justification, deontological justification, all $p < .001$). Additionally, participants judged an excuse more positive than a deontological justification ($p < .001$).

Study 1

Contrary to the results of the pilot study, the form of neutralization influenced both the perception of the protagonists power, $F(5, 348) = 9.90$, $p < .001$, $\eta^2 = .12$, and the judgment of the neutralization used, $F(3, 216) = 20.17$, $p < .001$, $\eta^2 = 21.89$. Post hoc tests showed that honest protagonists were perceived as significantly more powerful than protagonists using an apology, excuse, or a deontological justification. Only protagonists using a consequentialist justification did not differ from the honest condition in perceived power. Post hoc tests for the judgment of neutralization used showed that an apology was judged more positively than all other neutralizations used (all $p <$ .001), and that an excuse was judged more positively than a deontological justification ($p = .02$).

(*Appendices continue*)

Study 2

In accordance to Study 1, the neutralization had an effect on power perception, $F(5, 392) = 4.17$, $p = .001$, $\eta^2 = .05$, and the judgment of the excuse, $F(3, 262) = 18.17$, $p < .001$, $\eta^2 = .17$. Post hoc tests showed that protagonists who were honest were perceived as more powerful than protagonists who used a consequentialist justification. For the judgment of the excuse, using an apology was judged more positively than using an excuse or justification. No other conditions differed significantly from each other.

Appendix D

Descriptives of Manipulation Checks Across Studies

Table D1

Means and Standard Deviations (in Brackets) of Manipulation Checks for Each Condition (Pilot)

	Manipulation check variable			
Condition	Felt sorry	Used excuse	Justified means	Had the right
Apologizing	5.24 (1.26)	3.37 (1.40)	3.26 (1.37)	2.74 (1.35)
Excuse-making	4.74 (1.62)	4.74 (1.62)	4.11 (1.50)	4.29(1.41)
Consequentialist justification	2.11 (1.35)	5.03 (1.75)	4.87 (1.34)	4.63 (1.65)
Deontological justification	2.23 (1.55)	5.44 (1.68)	4.79 (1.66)	5.15 (1.89)

Table D2

Means and Standard Deviations (in Brackets) of Manipulation Checks for Each Condition (Study 1)

	Manipulation check variable			
Condition	Felt sorry	Used excuse	Justified means	Had the right
Apologizing	5.25 (1.58)	2.30 (1.52)	3.18 (1.40)	2.26 (1.32)
Excuse-making	4.64 (1.61)	5.69 (1.56)	3.69 (1.58)	2.94 (1.55)
Consequentialist justification	2.95 (1.63)	5.71 (1.42)	5.76 (0.96)	5.80 (1.04)
Deontological justification	2.08 (1.15)	5.51 (1.69)	5.69 (1.17)	5.90 (1.37)

Table D3

Means and Standard Deviations (in Brackets) of Manipulation Checks for Each Condition (Study 2)

	Manipulation check variable			
Condition	Felt sorry	Used excuse	Justified means	Had the right
Apologizing	5.17 (1.61)	2.73 (1.61)	2.76 (1.53)	2.80 (1.49)
Excuse-making	3.87 (1.72)	5.79 (1.47)	4.18 (1.67)	3.77 (1.81)
Consequentialist justification	2.09 (1.52)	5.48 (1.63)	5.72 (1.51)	5.75 (1.49)
Deontological justification	1.75 (1.26)	5.51 (1.84)	5.29 (1.67)	5.97 (1.51)

Received January 24, 2022
Revision received March 17, 2023
Accepted April 24, 2023 ■