Righting a wrong: Retaliation on a voodoo doll symbolizing an abusive supervisor restores justice

Lindie H. Liang, Douglas J. Brown, Huiwen Lian, Samuel Hanig, D. Lance Ferris, Lisa M. Keeping

Abstract

When a subordinate receives abusive treatment from a supervisor, a natural response is to retaliate against the supervisor. Although retaliation is dysfunctional and should be discouraged, we examine the potential functional role retaliation plays in terms of alleviating the negative consequences of abusive supervision on subordinate justice perceptions. Based on the notion that retaliation following mistreatment can restore justice for victims, we propose a model whereby retaliation following abusive supervision alleviates the negative effect of abusive supervision on subordinate justice perceptions. In two experimental studies (Study 1 and 2), whereby we manipulated abusive supervision and subordinate symbolic retaliation—in particular, harming a voodoo doll that represents the abusive supervisor—we found general support for our predictions. Theoretical and practical implications are discussed.

Introduction

When a subordinate is subjected to abusive supervision such as public ridicule, yelling, scapegoating, or other forms of supervisor mistreatment, a natural response for the subordinate is to directly retaliate against the abusive supervisor (Bies & Tripp, 1996). Indeed, a growing body of studies (e.g., Lian, Brown, Ferris, Liang, Keeping, & Morrison, 2014; Mitchell & Ambrose, 2007) and meta-analyses (Mackey, Frieder, Brees, & Martinko, 2017; Schyns & Schilling, 2013) suggests that a relationship exists between abusive supervision and subsequent subordinate retaliation. Unfortunately, retaliation—or actions “in response to some perceived harm or wrongdoing by another party that is intended to inflict damage” (Aquino, Tripp, & Bies, 2001, p. 53)—would seem to have destructive consequences for all parties involved. For instance, retaliation is detrimental to supervisor-subordinate relationships, such that it can escalate conflict, resulting in further acts of supervisory abuse (Aquino et al., 2001; Pruitt & Rubin, 1986; Tepper et al., 2009). Moreover, retaliation can result in expensive lawsuits (Perry, 2000) as well as undermine employee job performance (Robinson & Greenberg, 1998). Given these negative effects, various researchers have argued that retaliation should be avoided (e.g., Folger & Baron, 1996; Lian, Brown, et al., 2014).

Yet, despite these negative consequences, retaliation appears to be relatively common. For example, surveys have shown that 76% of employees reported engaging in aggression towards their supervisor over the past year (Greenberg & Barling, 1999), and that employees aggress towards their supervisor as much as they do towards other coworkers, perhaps more so (Baron, Neuman, & Geddes, 1999). The prevalence of retaliation suggests that retaliation may play a functional role in dealing with abuse—a perspective largely overlooked in the abusive supervision literature (for an exception, see Tepper, Mitchell, Haggard, Kwan, & Park, 2015). In fact, numerous perspectives that support the notion argue retaliation exists as a phenomenon precisely because it can be adaptive. For example, a social functionalist perspective of behavior would argue that retaliation exists because it serves an adaptive response (Keltner & Gross, 1999); a rational actor perspective would argue that retaliation occurs because actors conclude it serves a purpose (Vroom, 1964); and a social exchange perspective...
Subordinate Retaliation

Abusive Supervision

Subordinate Injustice Perceptions

Fig. 1. Heuristic model.

would argue that retaliation occurs, because it helps restore balance in a relationship (Cropanzano & Mitchell, 2005). Nevertheless, empirical evidence demonstrating the adaptive or functional nature of retaliation for the retaliator is scant.

Drawing inspiration from these perspectives—as well as frameworks which regard abusive supervision as undermining justice perceptions (Tepper, 2000) and retaliation as a reaction to injustice (Skarlicki & Folger, 1997)—we propose a functional theory of retaliation whereby engaging in retaliation reaffirms one’s sense of justice (see Fig. 1). In this framework, abusive supervision acts as an external stressor (Liang, Hanig, Evans, Brown, & Lian, in press; Restubog, Scott, & Zagenczyk, 2011) that violates people’s expectations for fair treatment (e.g., Adams, 1965; Lerner, 1980).

In presenting our functional theory of retaliation, our work makes several important contributions to the literature. First, our work contributes to both the literature on justice frameworks of the consequences of abusive supervision, as well as the literature on retaliation in the workplace. Though it has been posited that abusive supervision is unfair (Tepper, 2000), and that retaliation in response to being wronged can serve as a means for individuals to restore justice (e.g., Bies & Tripp, 1998; Bies & Tripp, 2002; Greenberg, 1990; Schyns & Schilling, 2013; Tepper et al., 2009), prior abusive supervision research has typically only considered subordinate retaliation as an outcome of abusive supervision (for an exception, see Tepper et al., 2015). In the current research, we directly test retaliation as a means of restoring justice rather than simply as a response to perceived injustice, by considering the interactive effect of abusive supervision and retaliation on perceptions of justice.

Second, our work contributes to the retributive justice literature by examining the beneficial effects of retaliation on victim outcomes. The dominant perspective of the retaliation literature is that retaliation is principally destructive, and therefore places victims of mistreatment who retaliate in the wrong. However, our work adopts the perspective of the victim and puts forth a functional view that retaliation buffers the detrimental impact of abusive supervision by directly restoring the justice perceptions of victims. As such, our research “gives back” (Whetten, Felin, & King, 2009) to the retributive justice literature in which it is grounded and enriches the literature by providing a more nuanced understanding of the outcomes of retaliation.

Third, our work extends social exchange frameworks of retaliatory responses to abusive supervision (Mitchell & Ambrose, 2007; Thau & Mitchell, 2010). In particular, prior social exchange research has primarily focused on how supervisor behaviors lead to subordinate outcomes—presenting what is essentially a stimulus-response perspective of social exchange, whereby subordinate outcomes are determined by supervisor inputs (as noted by Tepper et al., 2015). In contrast, our paper adopts a relational perspective (Aquino & Lamertz, 2004), whereby subordinates actively engage in actions affecting the overall exchange, with such actions also impacting subordinates’ own outcomes (in particular, their sense of justice). In so doing, we respond to the call of Tepper et al. (2015) to consider both sides of the exchange relationship when examining social exchange outcomes.

Finally, our work contributes to the abusive supervision literature and the leadership literature in general. Much of the research on leadership styles and outcomes is premised almost entirely upon cross-sectional field studies, assuming that leadership style causes subordinate outcomes without any concrete evidence. Cross-sectional designs are not only limited in drawing conclusions (Fischer, Dietz, & Antonakis, 2017), but are also susceptible to potential endogeneity threats (Antonakis, 2017); thus, the interpretation of those findings is limited. Our work addresses those issues in the literature by advancing well-designed experimental paradigms; in so doing, we respond to the call for leadership scholars to go beyond the “cross-sectional snapshots” with more creative experimental designs in leadership research (Antonakis, 2017, p. 12; Brown & Lord, 1999).

Justice frameworks of abusive supervision

People care about justice and have a fundamental need to believe that we live in a world that is a fair and orderly place where individuals get what they deserve (Lerner, 1980). The concern for justice is universal and serves the evolutionary function of promoting long-term cooperation, which is critical to the survival of the human species (Brosnan & de Waal, 2003). Justice is also hedonically valued by human beings, as evidenced by functional magnetic resonance imaging studies showing that people’s brain regions associated with reward processes are activated when receiving fair rather than unfair monetary offers (Tabibnia, Satpute, & Lieberman, 2008).

This fundamental concern for justice is central to early theories of distributive and procedural justice (Leventhal, 1980; Thibaut & Walker, 1975), which suggest that people care about the fair allocation of resources and the use of fair procedures, because these outcomes serve instrumental purposes (Tyler, 1987). When fair reward distributions and procedures are in place, people believe that they will be rewarded and punished proportionately to their actions; as a result, people are encouraged to work hard towards their goals and refrain from harming others (Hafer, 2000; Hafer, Bègue, Choma, & Dempsey, 2005). In addition to instrumental reasons, people also care about justice because it communicates relational information; in particular, being treated fairly carries implications about people’s social standing in their group (Lind & Tyler, 1992). People infer their social standing in a group from the treatment they receive: fair treatment conveys positive social identity-relevant information for individuals and signifies that they are valued members of the group (Lind & Tyler, 1988; Tyler & Blader, 2003). On the other hand, being treated without dignity and respect is perceived to not only hurt a victim’s standing within the group (Mitchell, Vogel, & Folger, 2015; Tyler & Lind, 1992), but also damage the victim’s self-worth (Ferris, Spence, Brown, & Heller, 2012; Tepper, 2000). All of the above suggests that concern for justice is a powerful motivational force that drives behaviors (Cropanzano, Byrne, Bococel, & Rupp, 2001), and that people are motivated to see justice prevail and be reaffirmed in their belief that people get what they deserve (Colquitt, Greenberg, & Zapata-Phelan, 2005).

Drawing on the central role justice plays in our everyday lives, abusive supervision research has used justice frameworks to explain the detrimental effects of being abused (Tepper, 2000). Abusive supervision represents supervisory behaviors that are non-physical in nature but nonetheless convey a sense of hostility towards subordinates. Such behaviors typically include ridiculing and humiliating subordinates in public, refusing to speak with subordinates, or otherwise debasing subordinates (Tepper, 2000). Extensive research, albeit cross-sectional in nature, has established relationships between abusive supervisory behaviors and subordinates’ diminished justice perceptions (Aryee,
reduce the aversive state (Colquitt et al., 2005).

Justice restoration: the function of retaliation

One way to alleviate the aversive feelings of unfair treatment is by engaging in actions to get even, which can safeguard one's perceptions that the world is a just place. In particular, although unfair treatment thwarts people's sense of justice (Bobocel & Hafer, 2007), retaliating against the harm-doer restores a sense of justice and thus affirms people's perceptions that those who do bad get what they deserve (Lerner, 1998). Retaliation following a transgression is a widespread social expectation that is at once primitive, culturally universal, and an organizing principle in human affairs (Hogan & Emmer, 1981; Tyler & Boeckmann, 1997). To elaborate, if a person violates the norm of respect and mistreats another party, these actions may thwart the receiving party's belief in a non-random and just world (e.g., Kay, Whitson, Gaucher, & Galinsky, 2009; Lerner, 1980). Consequently, the receiving party will seek to bring order to the situation by engaging in retaliatory behaviors intended to make the harm-doer pay (Skarlicki & Folger, 1997). This idea echoes Adams' (1965) equity theory, which describes feelings of injustice as similar in nature to aversive dissonance and psychological tension; as such, the individual experiencing injustice is motivated to take action to restore justice, in an effort to reduce the aversive state (Colquitt et al., 2005).

Consistent with this line of reasoning, research has shown that people are hardwired to punish norm violators, even at the expense of incurring personal cost (de Quervain et al., 2004; Fehr & Gachter, 2002; Henrich et al., 2006). Retaliating against a norm violator activates the reward-related region of their brain; thus, people anticipate feelings of relief and satisfaction of “sweet revenge”. Moreover, research has shown that delinquent acts of retaliation allow for individuals to maintain a sense of justice in their social interactions (Brezina, 1996), which dovetails with the argument that employees retaliate with the purpose of restoring justice perceptions (Bies & Tripp, 1998; Scott, Colquitt, & Paddock, 2009; Tripp, Bies, & Aquino, 2002). Thus, retaliating against an offender in proportion to the harm done to the victim can be seen as restoring justice for the victim (Okimoto, Wenzel, & Feather, 2012).

The moderating role of retaliation

We suggest that engaging in retaliation should mitigate the negative impact of abusive supervision on subordinates’ justice perceptions. When a supervisor belittles and insults a subordinate, it violates the subordinate's need to maintain justice (Bobocel & Hafer, 2007; Lerner, 1980). This violation will create an unpleasant tension accompanied by the desire to reaffirm justice (Lerner, 1980), and this desire will remain activated until it is fulfilled (Denzler, Förster, & Liberman, 2009; Marsh, Hicks, & Bryan, 1999; Zeigarnik, 1938). In line with Adams’ equity theory, which posits that rewards and punishments should be distributed proportionate to what the other party deserves (Adams, 1963; Homans, 1974), the opportunity to act deviantly towards an abusive supervisor evens the score and serves as a means to fulfill the desire to reaffirm violated justice. Once justice has been restored through retaliation, subordinates should be less likely to perceive injustice (Denzler et al., 2009).

However, not all subordinates retaliate following abusive supervision. Some subordinates may withhold retaliation because direct retaliation against their supervisor may not always be possible (Dollard, Miller, Doob, Mowrer, & Sears, 1939), or there may be formal organizational mechanisms in place to prevent subordinates from engaging in retaliation (Bies & Tripp, 1998). When subordinates are unable to retaliate against an abusive supervisor, the desire to reaffirm justice will remain activated and subordinates’ injustice perceptions linger (Denzler et al., 2009). In line with this argument, it has been suggested that thwarted retaliation may not only leave subordinates ruminating about the unfair event (Bies & Tripp, 1996; Bies & Tripp, 1998), but also perpetuate their sense of unfulfilled justice (Carlsmith, Darley, & Robinson, 2002). In the context of abusive supervision, we suggest that retaliating against an abusive supervisor will mitigate the positive effect of abusive supervision on subordinates' injustice perceptions.

Hypothesis 1. Retaliation moderates the positive relation between abusive supervision and subordinate injustice perceptions, such that the relationship is weaker when retaliation is high rather than low.

Overview of studies

A main challenge in studying abusive supervision in the field setting is that it is a low base-rate phenomenon, thus the mean and variance of abusive supervision are restricted. It is possible that the low base rate of abusive supervision could be due to selection effects (Breevaart & de Vries, 2017), such that participants who are comfortable rating their supervisors tend not to perceive their supervisors as abusive. Moreover, it has been suggested that given the highly undesirable behaviors associated with the abusive supervision construct, the low variance could be due to response style bias in that items that reflect different degrees of abusive supervision are answered in a uniform fashion (Breevaart & de Vries, 2017). Finally, as with any field studies, there is a potential endogeneity threat (Antonakis, 2017) in that the variance in subordinates’ perceptions of abusive supervision could be explained by something that is not modelled (e.g., subordinate personality characteristics, Wu & Hu, 2009). Given these challenges in studying abusive supervision in the field settings, as well as to complement the existing abusive supervision literature that has been based largely on field studies, we present a series of experimental studies with valid, representative, and strong manipulations to test our theory (Highhouse, 2009).

The purpose of these studies is twofold. First, to reduce any risk of demand characteristics associated with experimental methodology, we used an implicit measure to capture our dependent variable—subordinate injustice perceptions—which allows us to assess injustice perceptions without participants being aware of what was being assessed (Uhlmann et al., 2012). Second, because the abusive supervision literature has relied almost exclusively on field studies, we sought to establish greater causal confidence in our predicted model by experimentally manipulating both abusive supervision and retaliation. We test Hypothesis 1 in two experimental studies (Study 1 and 2), whereby we manipulated abusive supervision and retaliation, while implicitly assessing participants' justice perceptions.

Pilot study: injustice word fragment development

We used a word fragment measure to assess implicit injustice perceptions. Implicit perceptions of injustice involve the unconscious activation of injustice-related concepts in the minds of individuals (Smith & DeCoster, 2000; Strack & Deutsch, 2004), which enhances the extent to which these injustice related concepts are accessible to individuals (Johnson & Saboe, 2011; Kunda & Thagard, 1996; Smith, 1996; Uhlmann et al., 2012). Such a word fragment approach is a reliable and valid method for assessing implicit cognitive processes (Vargas, Sekaquaptewa, & von Hippel, 2007), and past studies have developed and used the word fragment method to assess implicit activation of various constructs, such as aggression and self-concept (e.g., Anderson, Carnegie, & Eubanks, 2003; Johnson & Lord, 2010).
Participants were provided with five word fragments and were asked to complete the fragments with the first word that comes to mind and would both complete the fragments and form a meaningful word. Each of the five word fragments can be completed to form either an injustice related word or a neutral word. For example, the fragment un_eual can be completed as “unusual” (i.e., neutral word), or “unequal” (i.e., injustice word). The ratio of injustice to total words that participants completed provides a measure of implicit injustice perceptions\(^2\) (see Gilbert & Hixon, 1991; Johnson & Saboe, 2011). We developed and pilot tested these word fragments with the procedures described below.

We first identified five injustice-related target words (wrong, unfair, unequal, unjust, and violated) from a published study that measured implicit justice (Hafer, 2000). Next, we developed word fragments for each target word using the English Lexicon Project database (Washington University, 2009), such that each word fragment can be completed as either a target (i.e., injustice-related) word, or a neutral word. Specific word fragments were chosen to ensure that the word frequency of the target word was approximately equal to the word frequency of the neutral word. Because multiple variations of word fragments exist (e.g., the target word “unequal” can have a fragment “un_ e _al,” which yields a matching word frequency neutral word “aseexual,” or a fragment “un _ eal,” which yields a matching word frequency neutral word “unusual”), we pilot tested our implicit injustice measure before the main study. We wanted to ensure we chose the fragments that (1) yield an optimal target/neutral words balance (i.e., the fragments were not biased towards a target or neutral words in the absence of an experimental manipulation); (2) yield the least amount of blanks (i.e., the fragment is not too difficult to complete); and (3) the neutral words were indeed neutral and not perceived as being related to justice.

To validate the injustice word fragments, we recruited 204 native English speakers (40% male; Age: M = 35.14, SD = 12.16) from Amazon’s Mechanical Turk (Mturk) who were compensated $0.25 in exchange for their participation. Following the procedure by Johnson and Saboe (2011), participants were administered an online survey that contained all implicit word fragments, and were instructed to complete the fragments as quickly as possible and skip any word fragments if no word came to mind immediately. Based on pilot test results, we retained the final five word fragments that are considered to have adequate variance in responses (i.e., at least 5% and no > 80% of the entries were the target word), and are relatively easy to complete (i.e., over 60% of participants did not leave the fragment blank). Finally, to assess whether any of the neutral words represent the construct of injustice, we provide a definition of the construct injustice for participants and asked them to rate the extent to which the neutral words represent injustice on a scale of 1 (not at all representative) to 7 (extremely representative). None of the neutral words for the final retained word fragments are especially representative of the construct injustice (Range: 2.01–2.67; M = 2.32, SD = 0.24). The final five word fragments used in this study are included in Appendix A.

**Study 1 method**

**Participants**

We recruited our participants from Amazon’s Mturk. If proper safeguards are taken (see Meade & Craig, 2012), participants recruited from Mturk tend to yield high quality data, tend to be motivated to complete tasks even without any financial incentives (Buhmester, Kwang, & Gosling, 2011), and tend to be more attentive than participants drawn from undergraduate student subject pools (Hauser & Schwarz, 2015).

We advertised our study on Mturk to 229 full time employees living in the USA or Canada who had experience working with a supervisor. In exchange, we deposited $1.00 payment to participants’ Mturk accounts. Following the recommendations by Meade and Craig (2012), we excluded 2 participants who reported post-experiment that they did not wish to have their data included in the study, and 32 additional participants who reported that they did not engage in the experimental manipulations used in this study. This left us with a final sample of 195 participants (46% male; Age: M = 35 years, SD = 9.97).

**Procedure**

In a between-subjects design, participants were randomly assigned into three conditions: abusive supervision/no retaliation, abusive supervision/retaliation, and a control condition. In the experimental conditions (i.e., abusive supervision/retaliation, and abusive supervision/no retaliation), participants were first asked to recall and visualize a workplace interaction, which was used to induce abusive supervision. Following this task, participants were asked to work on another task that involved the use of an online voodoo doll for 1 min, which we used to manipulate retaliation against the supervisor. Participants were then asked to work on another ostensibly unrelated task that involved completing five word fragments, which was used to assess participants’ implicit injustice perceptions. In the control condition, participants were simply asked to solve the word fragments. Finally, participants in all conditions completed a brief demographics survey.

**Inducing abusive supervision**

To induce abusive supervision, we used the critical incident technique (e.g., Liang, Lian, Brown, Ferris, Hanig, & Keeping, 2016; Liang, Brown, Ferris, Hanig, & Keeping, in press). Participants were first given a definition of a supervisor: “A supervisor is the individual that you report directly to, or who is responsible for assessments of your work” and were asked to visualize their supervisor. Consistent with the construct of abusive supervision (Tepper, 2000), participants were then asked to recall and visualize an incident in which their supervisor treated them with hostile verbal and/or nonverbal behaviors, such as being rude to them, making negative comments about them, and failing to acknowledge their hard work. Considering abusive supervision is a low base-rate phenomenon (Tepper et al., 2017), we did not specify a timeframe within which the incident occurred, thus increasing the likelihood that participants will recall such an event that has happened to them throughout their lifespan.

**Manipulating retaliation**

We used a voodoo doll task (VDT) paradigm to manipulate retaliation. The VDT is a validated task that has been used for measuring harming behaviors (DeWall et al., 2013; Finkel et al., 2012; Slotter et al., 2012). This task usually involves participants stabbing a doll that represents a specific person (e.g., spouse, offspring, reviewer, or someone who has offended them; Bushman, DeWall, Pond, & Hanus, 2014; McCarthy, Crouch, Basham, Milner, & Skowronski, 2016; Denzler et al., 2009; Liang, Brown, et al., in press), and the goal of the task is to symbolically harm the doll that represents the specific person. The underlying mechanism of the VDT is based on the law of similarity (Rozin, Millman, & Nemeroff, 1986) whereby people project characteristics of the person onto the voodoo doll; thus, the process of harming the voodoo doll bears psychological similarities to the process of causing harm to the person that the voodoo doll symbolically represents (DeWall et al., 2013).

In our study, we manipulated retaliation by instructing participants to harm a voodoo doll that represents their supervisor. In particular, in the abusive supervision/retaliation condition, we instructed...
participants first go to a website (http://www.dumb.com/vooodoodoll) where they encountered an online voodoo doll. They were then asked to label the voodoo doll with their supervisor's initials. Next, we asked the participants to use the materials provided (e.g., pins, pliers, fire) on the doll over the next minute. In the abusive supervision/no retaliation condition, participants were shown a screenshot of the voodoo doll from the website, they were asked to label the doll as "Nobody", and trace the outline of the doll with a cursor over the next minute.

**Manipulation checks**

**Abusive supervision manipulation check**

Prior to collecting Study 1 and 2 data, we pilot tested the abusive supervision manipulation with an independent sample of 352 students from a mid-sized university. We used the 5-item short abusive supervision measure (Mitchell & Ambrose, 2007; Tepper, 2000) to check for our abusive supervision manipulation. Following the abusive supervision manipulation, participants rated the extent to which they agree with each statement based on their thoughts and feelings about their visualized supervisor on a 5-point Likert scale (1 = I can't remember him/her ever using this behavior with me; 5 = he/she uses this behavior very often with me). Sample items include “My visualized supervisor ridicules me,” and “My visualized supervisor tells me my thoughts or feelings are stupid” (α = 0.95). There was a significant effect of the abusive supervision manipulation on participants’ ratings of perceived abusive supervision of the visualized supervisor (t(350) = -8.92, p < .01, d = -0.95), such that participants in the abusive supervision condition perceived greater abusive supervision of the visualized supervisor (M = 2.77, SD = 1.19) compared to those in the neutral interaction condition (M = 1.73, SD = 1.00). This indicates that our manipulation of abusive supervision was successful.

**Retaliation treatment check**

To ensure participants in the treatment conditions attended to the treatments (i.e., those in the retaliation condition harmed the voodoo doll representing their supervisor, and those in the no retaliation condition did not perceive the doll as their supervisor and did not harm the doll in any way), we administered treatment checks immediately following the VDT by directly asking the participants to recall what they were asked to do during the experiment. Participants who did not attend to the treatment should be excluded (Sigall & Mills, 1998), as participants’ failure to attend to treatment is akin to having an equipment failure, rendering the data meaningless. We asked the participants in the abusive supervision/retaliation condition: “In the previous task, did you use any of the materials (pins, pliers, or fire) on the doll?” Participants who self-reported that they did not harm the doll (n = 9) were excluded from the analyses. We asked participants in the abusive supervision/no retaliation condition: “In the previous task, who did you imagine the voodoo doll as representing?” Participants who self-reported that the doll represented someone other than ‘Nobody’ as instructed (n = 24) were excluded from the analyses.

**Study 1 results**

**Hypothesis testing**

Table 1 presents the means, standard deviations, and correlations of the variables for Study 1. To test the hypothesis that not retaliating following abusive supervision engenders greater perceptions of injustice compared to having the opportunity to retaliate (Hypothesis 1), we conducted a One-Way Analysis of Variance (ANOVA). In support of Hypothesis 1, there was a significant effect of condition (i.e., control condition, abusive supervision/retaliation, and abusive supervision/no retaliation) on participants’ implicit injustice perceptions [F(2, 192) = 3.81, p = .02, η² = 0.04]. As illustrated in Fig. 2, planned contrasts revealed that participants who did not engage in retaliation (M = 0.27, SD = 0.24) experienced significantly higher injustice perceptions compared to participants who engaged in retaliation (M = 0.19, SD = 0.22, t(192) = -2.19, p = .03, r = 0.16), and participants in the control condition [M = 0.17, SD = 0.19, t (192) = -2.64, p = .01, r = 0.19]. Moreover, participants in the abusive supervision/retaliation condition did not significantly differ from participants in the control condition in terms of their implicit injustice perceptions [t(192) = -0.48, p = .63, r = 0.03].

---

3 While, in the current study, we did not have a manipulation check in the traditional sense to assess whether the retaliation manipulation was perceived psychologically by participants, future studies could overcome such a limitation by administering retaliation manipulation check items that assess the extent to which participants attempted revenge against their supervisors (e.g., to what extent you tried to get even with your supervisor? and “to what extent you tried to make something happen to your supervisor?”). These items are adapted from Aquino et al. (2001) and can be used after assessing the dependent variable.

4 Recoding those participants as “abusive supervision/no supervisor-directed deviance” did not change the pattern of the results.

5 Including the participants who failed the compliance check in the analyses did not change the pattern of the results.
Robustness checks

Robustness checks for endogeneity

Given that there are participants who did not comply with our experimental manipulations (i.e., $n = 9$ in the abusive supervision/retaliation condition and $n = 24$ in the abusive supervision/no retaliation condition), there may be a selection problem in terms of whether participants decided to comply or not. This selection problem may bias coefficients if not correctly modeled (Angrist & Imbens, 1995; Antonakis, Bendahan, Jacquart, & Lalive, 2010), because participants who decided not to attend to our manipulation may differ from participants who attended to our manipulation. In other words, even though we experimentally manipulated retaliation, retaliation may not be exogenous in that it might be correlated with omitted causes (Antonakis et al., 2010). In order to model this selection correctly, we estimated the local average treatment effect (LATE), as recommended by Antonakis et al. (2010). By estimating the LATE, we take into account that some participants in our experimental conditions decided not to attend to the manipulation; thus, compared to our original analyses, in which we estimated the average treatment effect (ATE) by excluding the non-compliers, the LATE estimates represent an unbiased and more accurate coefficient of the effect of our manipulation on the dependent variable. The LATE was estimated with structural equation modeling approach using Mplus. Using maximum likelihood estimation, we explicitly modelled the selection in terms of whether participants complied with the treatment in the retaliation condition; we used an instrumental-variable estimator by modeling the covariance between the dependent variable and whether the participant complied (which is an endogenous choice). That is, in the equations below, $c_1$ denotes the complier variable in the abusive supervision/retaliation condition and $c_2$ denotes the complier variable in the abusive supervision/no retaliation condition. For $c_1$ and $c_2$, non-compliers were coded as 0, and compliers were coded as 1. Moreover, $x_{d1}$ denotes the dummy instrumental variable of abusive supervision/retaliation condition, and $x_{d2}$ denotes the dummy instrumental variable of abusive supervision/no retaliation condition. The baseline condition, the omitted category for the dummies, is the control condition. The error terms of $y$, $c_1$, and $c_2$ were specified to covary.

$$
\begin{align*}
&c_1 = b_0 + b_1 x_{d1} + b_2 x_{d2} + \epsilon_1 \\
&c_2 = b_0 + b_1 x_{d1} + b_2 x_{d2} + \epsilon_2 \\
&y = g_0 + g_1 c_1 + g_2 c_2 + \epsilon_3
\end{align*}
$$

Table 2 presents a comparison of three sets of results—ATE with non-compliers included (Panel A), ATE without non-compliers (Panel B), and LATE (Panel C). As shown in Table 2, the LATE estimates are close to the ATE estimates without the non-compliers (i.e., $+1.62\%$ difference in abusive supervision/retaliation condition, and $+5.17\%$ difference in abusive supervision/no retaliation condition); thus, it

---

Table 2

<table>
<thead>
<tr>
<th>Panel A</th>
<th>ATE with non-compliers</th>
<th>Predicted margin</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control condition</td>
<td>0.167</td>
<td>0.022</td>
<td></td>
</tr>
<tr>
<td>Abusive supervision/retaliation</td>
<td>0.186</td>
<td>0.025</td>
<td></td>
</tr>
<tr>
<td>Abusive supervision/no retaliation</td>
<td>0.249</td>
<td>0.014</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B</th>
<th>ATE without non-compliers</th>
<th>Predicted margin</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control condition</td>
<td>0.167</td>
<td>0.026</td>
<td></td>
</tr>
<tr>
<td>Abusive supervision/retaliation</td>
<td>0.185</td>
<td>0.026</td>
<td></td>
</tr>
<tr>
<td>Abusive supervision/no retaliation</td>
<td>0.271</td>
<td>0.030</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C</th>
<th>LATE</th>
<th>Predicted margin</th>
<th>Standard error</th>
<th>% difference in predicted margin between Panel B and C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control condition</td>
<td>0.167</td>
<td>0.025</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Abusive supervision/retaliation</td>
<td>0.188</td>
<td>0.027</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abusive supervision/no retaliation</td>
<td>0.285</td>
<td>0.037</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. $N = 227, 195, \text{and} 227$ for Panel A, B, and C, respectively. ATE = Average Treatment Effect, LATE = Local Average Treatment Effect.

---

Fig. 2. Condition in predicting implicit injustice perceptions (Study 1). Graphs were plotted using the predicted margins from Table 2 Panel B (i.e., ATE without non-compliers).
appears that there is no major selection problem in our data.

Robustness checks for implicit justice

Following prior studies (e.g., Johnson & Saboe, 2011), our dependent variable—implicit injustice perceptions—is computed as the ratio of injustice words to total words that participants completed, to eliminate the possibility that participants generate more words overall thereby inflating the number of injustice words (Johnson & Saboe, 2011). Given the known issues—such as biased estimates of the true relationship between variables—associated with using ratio variables as noted by researchers (Firebaugh & Gibbs, 1985; Kronml, 1993; Long, 1980), caution should be taken when using ratio variables. As such, we conducted supplementary analyses using the absolute number of injustice words as the dependent variable. Using ordinary least squares (OLS) estimation, one-way ANOVA results were consistent with the results we report using implicit injustice ratios. In particular, there was a significant effect of condition on participants’ implicit injustice perceptions \(F(2, 192) = 5.57, p = .004, \eta^2 = 0.05\). We also estimated the model with regression analysis in Mplus and obtained the exact same results \(F(2, 192) = 5.58, p = .004\).\(^6\)

Given that the number of injustice words (Mean = 0.78, SD = 0.91) contain many “zero” values, and that the standard deviation is higher than the mean, we re-estimated the model using the maximal likelihood estimation with robust standard errors (MLR) estimator in Mplus. The MLR estimator in Mplus uses a “sandwich” estimate of variance for cases to correct for where the nonnormality assumption of the variables is violated (Muthén & Muthén, 2015, p. 9). Using the MLR estimator, we obtained similar results as in our main analysis \(F(2, 192) = 4.46, p = .01\). Moreover, because the number of injustice words is a count variable, we repeated the above analysis with Poisson regression, which models count data. We estimated a zero-inflated Poisson model with the MLR estimator in Mplus. Consistent with our main analysis, condition significantly predicted the number of injustice words \(F(2, 192) = 5.57, p = .004\). To adjust for the over-dispersed count data (i.e., the standard deviation higher than the mean for injustice words), we also estimated a negative binomial model in Mplus, and we obtained results similar to those above \(F(2, 192) = 5.56, p = .004\). Finally, given that the number of injustice words can be also treated as an ordered categorical variable, we estimated a bounded count model using ordered probit regression with robust weighted least squares (WLSMV) estimator in Mplus. Ordered probit regression results were largely consistent with our main analysis \(F(2, 192) = 4.79, p = .008\).

Study 1 discussion

The above results indicate that, following abusive supervision, injustice perceptions are mitigated when participants retaliate (i.e., individuals in this condition did not differ from those who were in the control condition). To constructively replicate (Lykken, 1968) Study 1, we implemented a 2 × 2 experimental design in Study 2, whereby we manipulated levels of supervisory treatment (i.e., abusive supervision vs. neutral interaction) and levels of retaliation (i.e., no retaliation vs. retaliation). By experimentally manipulating both the independent and moderator variables, we provide a more fine-grained test of the justice restoration model and thus gain greater confidence in the validity of the hypothesis (Lykken, 1968).

Study 2 method

Participants

We advertised our study to 206 business school students from a mid-sized university who had previous work experience with a supervisor. In exchange, participants were given 0.5 course credit towards their classes. Following the recommendations by Meade and Craig (2012), we excluded 36 participants who reported post-experiment that they did not wish to have their data included in the study, and 24 participants who reported that they did not engage in the experimental manipulations used in this study (4 of the 24 participants also did not wish to have their data included in the study). This left us with a final sample of 150 students (57% male; Age: \(M = 19\) years, \(SD = 1.11\)).

Procedure

The procedure was largely similar to Study 1, except that we used a 2 (supervisory treatment: abusive supervision vs. neutral interaction) × 2 (retaliation: no retaliation vs. retaliation) between-subjects study design.

Manipulating supervisory treatment

We followed the same procedure as Study 1 to manipulate abusive supervision. In this study, a neutral interaction condition was included where participants were asked to recall and visualize a neutral interaction with their supervisor.
Retaliation treatment check

As in Study 1, we directly asked the participants in the retaliation condition to recall what happened during the experiment: “In the previous task, did you use any of the materials (pins, pliers, or fire) on the doll?” Participants who self-reported that they did not harm the doll (n = 24) were excluded from the analyses.7

Study 2 results

Hypothesis testing

Table 3 presents the means, standard deviations, and correlations of the variables for Study 2. To test the hypothesis that retaliation mitigates the effects of abusive supervision on injustice perceptions (Hypothesis 1), we conducted a 2 (supervisory treatment: abusive supervision vs. neutral interaction) × 2 (retaliation: no retaliation vs. retaliation) factorial ANOVA. Results revealed that there was a marginally significant Supervisory treatment × Retaliation interaction in predicting implicit injustice perceptions, \[ F(1, 146) = 2.90, p = .09, \] partial \( \eta^2 = 0.02 \). However, under the neutral interaction condition, the difference between the no retaliation condition (M = 0.16, SD = 0.23) and the retaliation condition (M = 0.21, SD = 0.23) in predicting implicit injustice perceptions was not significant \( [F(1, 146) = 0.52, p = .47, \text{partial } \eta^2 = 0.004]\).

Robustness checks

Robustness checks for endogeneity

As in Study 1, to ensure that the independent variable that we manipulated is truly exogenous and not influenced by participants’ self-selection to engage in the manipulation, we estimated the LATE with a structural equation modeling approach in Mplus. As with the first study, we used an instrumental-variable maximum likelihood estimation method to model the non-compliance selection by modeling the covariance between the dependent variable and the compliance variables.

In the equations below, \( c_1 \) denotes the compliance variable in the neutral/retaliation condition, and \( c_2 \) denotes the compliance variable in the abusive supervision/retaliation condition. For \( c_1 \) and \( c_2 \), non-compliers were coded as 0, and compliers were coded as 1. Moreover, \( x_{d1} \) denotes the dummy instrumental variable of the neutral interaction/retaliation condition, and \( x_{d2} \) denotes the dummy instrumental variable of the abusive supervision/retaliation condition, and \( x_{d3} \) denotes the dummy instrumental variable of the neutral interaction/no retaliation condition. The baseline (omitted) category is the abusive supervision/no retaliation condition. The error terms of \( y \), \( c_1 \), and \( c_2 \) were specified to covary.

\[ c_1 = b_0 + b_{1x}d_1 + b_{2x}d_2 + b_{3x}d_3 + e_1 \]

Table 4

Predicted margins of experimental conditions in predicting implicit injustice.

<table>
<thead>
<tr>
<th>Panel A</th>
<th>Panel B</th>
<th>Panel C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATE with non-compliers</td>
<td>ATE without non-compliers</td>
<td>LATE</td>
</tr>
<tr>
<td>Predicted margin</td>
<td>Standard error</td>
<td>Predicted margin</td>
</tr>
<tr>
<td>Neural interaction/no retaliation</td>
<td>0.163</td>
<td>0.033</td>
</tr>
<tr>
<td>Neutral interaction/retaliation</td>
<td>0.194</td>
<td>0.037</td>
</tr>
<tr>
<td>Abusive supervision/no retaliation</td>
<td>0.286</td>
<td>0.045</td>
</tr>
<tr>
<td>Abusive supervision/retaliation</td>
<td>0.224</td>
<td>0.036</td>
</tr>
</tbody>
</table>

Note. \( N = 170, 150, \) and \( 170 \) for Panel A, B, and C, respectively. ATE = Average Treatment Effect, LATE = Local Average Treatment Effect.

7 Recoding those participants as “no retaliation” did not change the pattern of the results.

Fig. 3. Abusive supervision x Retaliation in predicting implicit injustice perceptions (Study 2). Graphs were plotted using the predicted margins from Table 4 Panel B (i.e., ATE without non-compliers).
c2 = b0 + b1xd1 + b2xd2 + b3xd3 + e2
y = g0 + g1c1 + g2c2 + g3c3d + e3

Table 4 presents a comparison of three sets of results—ATE with non-compliers included (Panel A), ATE without non-compliers (Panel B), and LATE (Panel C). As shown in Table 4, the LATE estimates are fairly close to the ATE estimates without the non-compliers (i.e., 0% difference in the neutral interaction/no retaliation condition, −22.82% difference in the neutral interaction/retaliation condition, and +10.58% difference in the abusive supervision/retaliation condition); thus, it is possible that there is no significant selection problem in our data.

Robustness checks for implicit justice

As in Study 1, we conducted supplementary analyses using the absolute number of injustice words as the dependent variable. Using OLS estimation as we did in our main analyses, there was a marginally significant Supervisory treatment x Retaliation interaction in predicting implicit injustice perceptions (R2, 146 = 3.06, p = .08, partial η² = 0.02), replicating our results using implicit injustice ratio as the dependent variable. We also estimated the model with regression analysis in Mplus and obtained the exact same results [b = −0.49, SE = 0.27, p = .08, 95% CI (−0.994, 0.077)]. As in Study 1, injustice words (Mean = 0.67, SD = 0.84) contain many ‘zero’ values, and the standard deviation is higher than the mean. We thus re-estimated the model with robust standard errors using the MLR estimator in Mplus, and obtained similar results for the interaction term as in our main analysis [b = −0.49, SE = 0.28, p = .08, 95% CI (−1.03, 0.056)]. Moreover, we estimated a zero-inflated Poisson model with the MLR estimator to model the number of injustice words as a count variable. Consistent with our main analysis, the interaction term is marginally significant [b = −0.70, SE = 0.41, p = .09, 95% CI (−1.50, 1.044)]. To adjust for the over-dispersed count data (i.e., that the standard deviation is higher than the mean for injustice words), we also estimated a negative binomial model, and we obtained similar results for the interaction term as above [b = −0.70, SE = 0.41, p = .09, 95% CI (−1.50, 0.098)]. Finally, given that the number of injustice words can also be treated as an ordered categorical variable, we estimated a bounded count model using ordered probit regression with the WLSMV estimator, and the interaction term remained largely similar as in our main analysis [b = −0.67, SE = 0.39, p = .09, 95% CI (−1.42, 0.093)].

Probing ordinal interaction

As shown in Fig. 3, our test of the two-way interactive effect yielded evidence of an ordinal interaction (i.e., only one cell appears to differ from the others). It has been well-documented that such ordinal interactions are extremely difficult to detect (Bobko, 1986; Elias, 2004); as such, previous work (e.g., Oldham, Kulik, & Stepina, 1991; Spencer & Rupp, 2009) has typically followed the procedures outlined by Bobko (1986) to probe for ordinal interactions. Given the nature of our results, we opted to follow Bobko’s (1986) advice and probed the relationship between our four cells by conducting two contrasts: (1) assessing whether the three similar cell means are significantly different from each other; and (2) assessing whether the one different cell mean (i.e., abusive supervision/no retaliation) is significantly different from the average of other three cell means.

To test the first contrast, we conducted a One-Way ANOVA, which directly compared the neutral interaction/no retaliation, neutral interaction/retaliation, and abusive supervision/retaliation cells. As expected, the results of this contrast indicated no significant effect [R2, (110) = 0.31, p = .74, η² = 0.01]. Furthermore, subsequent follow-up tests in which we directly contrasted each possible pair of cells indicates no significant differences between the cells: neutral interaction/no retaliation (M = 0.16, SD = 0.23) was not significantly different from neutral interaction/retaliation (M = 0.21, SD = 0.23; t(110) = −0.75, p = .45, r = 0.07), neutral interaction/no retaliation was not significantly different from abusive supervision/retaliation (M = 0.19, SD = 0.25; t(110) = −0.51, p = .61, r = 0.05), and neutral/retaliation was not significantly different from abusive supervision/retaliation (t(110) = 0.30, p = .77, r = 0.03). Importantly, these results suggest that participants who were able to retaliate directly against their abusive supervisor did not report implicit injustice perceptions that differed significantly from participants in the two neutral interaction cells.

Next, we directly contrasted the abusive supervision/no retaliation cell against the average of the other three cell-means (i.e., neutral interaction/no retaliation, neutral interaction/retaliation, and abusive supervision/retaliation). The results of this contrast indicated that the participants in the abusive supervision/no retaliation condition (M = 0.29, SD = 0.28) had implicit injustice perceptions that were significantly higher than the average of the other three cells (t(146) = −2.11, p = .04, r = 0.17). The overall pattern of findings provides support for Hypothesis 1, in that implicit injustice perceptions were highest when an individual was unable to retaliate and were mitigated following acts of retaliation (i.e., these individuals did not differ from those who visualized a neutral interaction with their supervisor).

Study 2 discussion

Overall, consistent with Study 1, Study 2 results demonstrate that participants who recalled abusive supervisory treatment and were not given the opportunity to symbolically retaliate against the supervisor in question exhibited significantly higher implicit injustice perceptions compared to participants in the other conditions. In contrast, engaging in symbolic retaliation reduced the extent to which recalling abusive supervision led to implicit injustice perceptions.

General discussion

Although the abusive supervision literature has typically presented subordinate retaliation as dysfunctional reactions that should be discouraged, in the current paper, we offer a counter narrative by outlining a beneficial side of retaliation. In particular, we develop and test a theory that retaliation in response to abusive supervision restores justice perceptions. Across two experimental studies, we found strong empirical support for our hypothesis. Collectively, our results indicate that retaliation serves the function of restoring justice for victims of abusive supervision.

Theoretical implications

The extant literature largely tends to regard retaliation in response to abusive supervision as dysfunctional and self-defeating. Specifically, researchers have suggested that aversive outcomes such as lost rewards, punishment, or escalation of abuse may follow if subordinates directly confront an abusive supervisor (Aquino, Tripp, & Bies, 2006; Tepper et al., 2009). Given the possibility of such aversive outcomes, retaliation against an abusive supervisor has been regarded as an irrational behavior that results from a lack of self-control among abused subordinates (Lian, Brown, et al., 2014; Thau & Mitchell, 2010). However, the justice literature suggests that individuals have a fundamental motivation to maintain a sense of justice, thus, retaliation has been proposed as a way to restore justice when experiencing unfair treatment such as abusive supervision (Bies & Tripp, 1998). Building on this perspective, our work examines the benefits of retaliation for employees with respect to restoring justice for abused employees. Hence,
our work adds to the current understanding of retaliation and suggests that employees may not always engage in self-defeating responses towards abusive supervision (Tepper et al., 2015); rather, these responses serve to restore perceptions of justice for employees.

Our work contributes to the justice literature by directly testing the frequently posited assumption that retaliation functions to restore justice (Bies & Tripp, 1998), which has mainly been investigated in qualitative studies (see Bies & Tripp, 1996; Tripp & Bies, 1997). With evidence from two experimental studies, our research provides compelling evidence for the justice restoration view of retaliation and highlights the largely overlooked benefit of retaliation from the victim’s perspective.

Our work also extends the growing body of research on abusive supervision by adopting a relational perspective in understanding how subordinates react to abusive supervision. Past work examining subordinate outcomes of abusive supervision tends to regard subordinates as passive recipients of their supervisors’ mistreatment. However, recent research suggests that subordinates can actively affect the treatment that they receive rather than being purely passive recipients (Lian, Ferris, Morrison, & Brown, 2014). In line with the notion that harmful workplace behaviors such as abusive supervision should be understood within the context of a dyadic relationship (Aquino & Lamertz, 2004; Hershcovis & Barling, 2007; Hershcovis & Rafferty, 2012), our work takes into account both supervisor and subordinate behavior in determining the consequences of abusive supervision. In taking this approach, our work suggests that subordinates’ outcomes are not solely determined by their supervisors’ treatment. Instead, the impact of supervisor treatment on subordinates’ psyches is affected by subordinates’ own behavior.

Finally, our work takes a step towards advancing experimental methodologies in the study of abusive supervision to complement the existing abusive supervision literature that has been predominated by field studies. Given the inherent challenges in studying abusive supervision using field studies, and the call for leadership scholars to balance field studies with experimental designs (Antonakis, 2017; Brown & Lord, 1999), we designed a series of experimental studies to test a well-held theoretical assumption, thus providing a starting point for researchers to study abusive supervision using experimental methods. There are several advantages associated with our use of experimental methods to study abusive supervision. First, the experimental designs allow us to draw, with confidence, causal inferences regarding the finding that abusive supervision diminishes subordinate perceptions of justice, which has previously only been established with correlational research (Tepper, 2001, 2007). Second, experimental designs can get at implicit or nonconscious processes (Brown & Lord, 1999; Lord & Maher, 1991)—such as the implicit injustice perceptions that we assessed in Study 1 and 2—which are immune to demand characteristics and cannot be assessed with field studies. Third, experimental designs allow researchers to investigate constructs—such as abusive supervision and employee retaliation—which cannot be frequently observed or ethically manipulated in field settings. Finally, by manipulating the independent variables, experimental designs effectively reduce any potential endogeneity threats (Antonakis, 2017); thus, the results obtained from our study can be used to inform organizational practices and policy changes with greater confidence.

Despite the above benefits associated with experimental designs, organizational researchers continue to give priority to field studies (Highhouse, 2009). In particular, in the abusive supervision literature, one of the main concerns with experimentally studying the construct “has to do with issues of ecological validity” (Tepper et al., 2017, p. 131). Specifically, the construct involves “subordinates’ perceptions of the extent to which supervisors engage in the sustained display of hostile verbal and nonverbal behaviors, excluding physical contact” (Tepper, 2000, p. 178). Thus, the intricate leader-follower relationship that has developed over time is difficult to capture using experimental methodology. Below we offer recommendations for how experimental studies on abusive supervision can be conducted in a more robust manner.

To manipulate the construct abusive supervision in a controlled experimental setting—while simultaneously ensuring psychological realism to capture defining features of the construct—researchers could use a critical incident technique by asking participants to recall and visualize a workplace interaction with a supervisor in which their supervisor treated them with hostile verbal and/or nonverbal behaviors, such as being rude to them, making negative comments about them, and failing to acknowledge their hard work. The critical incident technique has been used in organizational research to elicit salient workplace experiences, such as the experience of workplace offences (e.g., Aquino et al., 2001; Aquino et al., 2006), supervisors’ experience of subordinate poor performance (Liang et al., 2016), and employees’ experience of hostility at work (Liang, Brown, et al., in press; Liang, Hanig, et al., in press).

We believe that this method can accurately capture the essential aspects of the construct of abusive supervision for the following reasons. First, by asking subordinates to recall a specific incident, the technique captures the “subordinates’ perception” aspect of the abusive supervision construct. Second, subordinates were asked to recall and visualize a real incident that happened to them, so this method does not rely on subordinates’ imagining being abused. Finally, visualizing an abusive supervisor could potentially trigger other incidents associated with this abusive supervisor. As such, this method is not simply a single shot exposure to a supervisor temporarily behaving in a hostile manner; rather, subordinates would have interacted with the recalled supervisor in real life, and their recalled supervisor could potentially have been treating them in a hostile manner over time. Thus, our methodology also captures the “sustained display” aspect of the abusive supervision construct.

Beyond using experimental designs to address the challenges in assessing abusive supervision in the field, we believe future research could also consider using the instrumental-variable design to estimate the effect abusive supervision on outcomes; that is, instrument variables such as leader personality, ability, or other stable individual differences could be used to remove possible endogeneity bias in ratings of abusive supervision to predict outcomes.

**Limitations and directions for future research**

As with all research, there are limitations to our work that need to be acknowledged and addressed in future studies. First, as theorized previously, abusive supervision represents a violation of the goal to maintain justice, and subordinates who experience abusive supervision are motivated to engage in actions to satisfy this goal. Given the equifinality property of goals, which refers to the idea that a goal can be achieved through multiple means (Austin & Vancouver, 1996), there should be multiple ways through which subordinates pursue justice. Retaliation is only one of the many means available. As a follow-up to the current study, a next step would be to simultaneously investigate the multiple means by which justice can be restored. For example, it has been suggested that, following a transgression, the act of forgiving the transgressor can free people from the inner turmoil that comes from harboring grudges and helps them to let go of any emotional injury that they have sustained (Adams, Zou, Inesi, & Pillutla, 2015, p. 131). Moreover, it has been suggested that acts of forgiveness, rather than bearing grudges, following a transgression can provide victims with important physiological as well as psychological benefits (Witvliet, Ludvig, & Laan, 2001; for a review, see Fehr, Gelfand, & Nag, 2010). Thus, it is possible that forgiveness of the transgressor is an alternative option that can bring closure and restore justice for the victim. Future
studies could be designed to compare and contrast retaliation with alternative means to bring closure and restore justice for the victim.

Second, in the current paper, we have only examined one function of retaliation as restoring justice for victims of abusive supervision. Yet, besides restoring justice, it has been proposed that retaliation serves a multitude of functions, which include deterring offenders from future mistreatment (Aquino et al., 2001; Tepper et al., 2009), serving a moral educative function by teaching the offender a lesson (Baumeister, 1997; Heider, 1958), saving face for the victim (Heider, 1958; McCullough, Bellah, Kilpatrick, & Johnson, 2001), increasing self-esteem, self-efficacy, improving the victims’ emotional state (Bies & Tripp, 1998), and, finally, maintaining the victims’ sense of control as well as restoring their sense of power and status (Katz, 1988; Okimoto et al., 2012). In this sense, our findings are limited in the extent to which they speak to the different theorized functions of retaliation; thus, future research could be directed to examine these different functions of retaliation.

Third, while the current research supports the beneficial effect of retaliation, it is plausible that retaliation may be harmful to the retaliator under certain circumstances, and that there may be individual differences in influencing the extent to which retaliation may or may not be beneficial to the retaliator. For instance, people who are high in moral identity (Aquino & Reed II, 2002) may feel that retaliation is inconsistent with their values and may thus feel worse following retaliation. On the other hand, people who are high in just world beliefs may feel a strong sense of vindication upon punishing an offender who is perceived to deserve such retribution (Kaiser, Brooke Vick, & Major, 2004). Future research should aim to examine these boundary conditions, given that they may influence whether retaliation is psychologically beneficial or harmful to the actor in question. Moreover, although the current research demonstrates that retaliation restores justice for victims of abusive supervision in the short term, the long-term impact of retaliation on the individual, organization, or society still remains unknown. To fully break the spiral of incivility (Anderson & Pearson, 1999), we first need to understand why employees engage in retaliation. The current research provides a test of theory that employees retaliate to restore justice, and, knowing this, future research needs to further investigate the long-term repercussions and devise effective interventions for employees to break out from the spiral of incivility.

Fourth, although in the current paper retaliation was performed directly by the victim, questions still remain as to whether merely observing a co-worker retaliate against an abusive supervisor could serve to restore justice for victims of abusive supervision, thus warranting future investigation. Moreover, although the current study addresses the beneficial effect of retaliation on victim well-being by restoring justice for victims, we have yet to address potential side effects of retaliation and justice restoration. In particular, while retaliation can help a victim of abusive supervision restore justice perceptions, it may also desensitize this individual’s justice perceptions. Thus, the suffering of other victims may be viewed with little sympathy and may not impel a desire to help (Bushman & Anderson, 2009). As such, future research could be directed at examining the unintended negative consequences of retaliation and justice restoration on third party reactions of witnessing abusive supervision.

Fifth, given that there are social costs associated with subordinates retaliating against their supervisor (e.g., lost financial rewards, and counterretaliation from the supervisor), a limitation of our retaliation manipulation is that it was not associated with any direct cost for subordinates. Future research could examine retaliation that is costly for subordinates. One way of doing so is by designing an altruistic punishment economic game whereby abused followers could choose to punish their abusive leader at a personal cost and without any personal benefit, but with the potential to benefit other players (Strobel, 2016) by deterring the leader from mistreating them in the future.

Finally, as with any experimental designs, despite the fact that our studies have strong internal validity, their ecological validity is limited in their ability to generalize to typical organizational settings. That being said, it has been suggested that strong experiments should “sacrifice real-world authenticity for internal validity” (Highhouse, 2009, p. 561). Moreover, given that each organization has its unique structure and culture, conducting field studies in one organization does not guarantee that it will generalize to other organizations (Brown & Lord, 1999; Highhouse, 2009). Thus, our goal was to design strong experiments that rigorously test our theoretical explanation, to ensure that our theory could generalize across organizations (Highhouse, 2009). Moreover, another limitation is that we asked participants to recall a time when they were abused by their supervisor; such method is known to trigger demand effects (e.g., Sturm & Antonakis, 2015), such that it is possible the effect of supervisory treatment on justice perceptions is due to participants changing their behaviors to conform to our hypotheses rather than due to our manipulations. However, such concern for demand effects should be minimized by using the implicit measure of injustice.

**Practical implications**

Although it is difficult to offer direct practical implications from our study, given that existing researches are suggestive that abusive supervision has a number of negative consequences (Tepper, 2007), our findings provide several indirect organizational implications for how some of these consequences may be alleviated. In particular, we have proposed and found that subordinate retaliation can directly influence subordinate justice perceptions. These findings suggest that retaliation not only benefits individual victims, but may also benefit the organization as a whole, given that justice perceptions is important for employee performance and well-being (Wright & Cropanzano, 1998; Wright, Cropanzano, & Bonett, 2007). Despite this, we do not mean to advise organizations to encourage or even tolerate subordinate retaliation, which can have significant organizational costs (Robinson & Greenberg, 1998); rather, efforts should be directed towards fostering subordinate justice perceptions. This can be achieved through implementing a zero-tolerance policy with regard to abusive supervision (e.g., Tepper, Duffy, Henle, & Lambert, 2006), thus preventing subordinates from experiencing violated justice perceptions in the first place and breaking the spiral of incivility (Andersson & Pearson, 1999). Moreover, given our findings that retaliation (i.e., by stabbing a voodoo doll meant to represent one’s supervisor) can deactivate injustice perceptions associated with abusive supervision, subordinates who receive perpetual mistreatment from their supervisor may benefit from harmlessly acts of symbolic retaliation against their supervisor.

More importantly, our studies suggest that organizations may want to take a new perspective on the deviant behavior of subordinates. In particular, rather than regarding subordinate retaliation as originating solely within the subordinate and punishing this subordinate as a result of their actions, organizations may want to look further into alternate cause of these behaviors and see whether they reflect larger organizational problems, such as unfair practices committed by management. Given that subordinates are likely to engage in retaliation as a way to restore their sense of justice, organizations may want to examine whether subordinate deviance is reflective of subordinate attempts to resolve perceived injustices, as well as reduce subordinate tendencies to resort to retaliation by addressing these issues.
Appendix A

Word fragment measure for implicit injustice used in Study 1 and 2.

<table>
<thead>
<tr>
<th>Word fragment</th>
<th>Target word</th>
<th>Neutral word</th>
</tr>
</thead>
<tbody>
<tr>
<td>U N _ _ U A L</td>
<td>Unequal</td>
<td>Unusual</td>
</tr>
<tr>
<td>_ _ O L A T E D</td>
<td>Violated</td>
<td>Isolated</td>
</tr>
<tr>
<td>R O N _</td>
<td>Wrong</td>
<td>Front</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Drone</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Crane</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Irony</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Irons</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prong</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prone</td>
</tr>
<tr>
<td>U N _ _ S T</td>
<td>Unjust</td>
<td>Unrest</td>
</tr>
<tr>
<td>U N _ A I _</td>
<td>Unfair</td>
<td>Unpaid</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unsaid</td>
</tr>
</tbody>
</table>

References


Aquino, K., Tripp, T. M., & Bies, R. J. (2001). How employees respond to personal o


Johnson, R. E., & Lord, R. G. (2001). Implicit e-


