**Abstract**

Individuals who harm others (i.e., transgressors) experience threat to their moral identity, which they often try to defend through moral disengagement (e.g., by denying their culpability). Optimistically, the needs-based model of reconciliation suggests that restoration of transgressors’ moral identity can reduce their defensiveness and increase their readiness for reconciliation. Several studies supported this possibility, revealing that morally accepting messages from their victims, as well as self-affirmation exercises through which transgressors affirmed their morality and the values breached by the transgression, increased their tendency to offer genuine, non-defensive apologies and invest effort in reconciliation. In contrast, however, morally accepting messages by third parties are associated with transgressors’ lower willingness to reconcile. Further research is needed to identify the conditions under which moral affirmation might lead to such ‘moral licensing’ effects, rather than to increased reconciliatory behavior. Understanding the effects of moral affirmation is practically important for structuring effective restorative justice procedures.

**Affirming transgressors’ morality as a strategy to promote apologies and interpersonal reconciliation: The promise and potential pitfalls**

Humans are a social animal (e.g., Dunbar, 2009), and conflict and transgressions are an inevitable part of social life. To enjoy the evolutionary benefits of sociality, humans, like other social primates (Silk, 2002), have evolved mechanisms for restoring valuable relationships in the aftermath of conflict (McCullough et al., 2013). In the seminal book *Mea culpa: A sociology of apology and reconciliation* sociologist Nicholas Tavuchis (1991) argues that apologies constitute the main mechanism to achieve this goal of harmony restoration, and that using them can dramatically transform the relations between former adversaries. Nevertheless, perpetrators of interpersonal transgressions often refrain from apologizing to their victim (Schumann, 2018).

Why is this the case, and what can be done about it? In the present chapter I will discuss the threat to transgressors’ moral identity as an obstacle to apologies and reconciliation and focus on moral affirmation as a means to increase transgressors’ readiness to apologize, compensate, and reconcile with their victims. Because research on moral affirmation is still in its early stages, I will also discuss the several unresolved issues that need to be studied in future research and explain why studying them has important practical implications.

**Obstacles to apology and reconciliation**

Interpersonal transgressions almost by definition involve transgressors’ violation of moral imperatives, such as fairness, loyalty, or prevention of harm. Such moral failures may occur because the transgressors hold low concern for the victims and the relationship with them, which is also a common reason for their reluctance to apologize (Schumann, 2018). However, even when transgressors do care about their victims, they may refrain from apologizing because they view their acts as justified or at least excusable under the circumstances (Schönbach, 1990). Indeed, there is a systematic ‘magnitude gap’ (a term coined by Baumeister et al., 1990) between victims and transgressors’ estimations of the immorality of the transgression and the severity of its consequences, which led Baumeister (1997) to conclude that “evil is in the eye of the beholder.” This may explain the counterintuitive finding that transgressors feel greater guilt for harm caused by unintentional rather than intentional acts (Schönbach, 1990) — although it is intentional rather than unintentional acts that meet the conditions for attribution of blame (controllability, foreseeability, and intentionality; Shaver, 1985). Whereas the unintentional harm transgressors have caused may catch them with their guards down, intentional acts — which were processed and thought about in advance — are more likely to be perceived by the transgressors as justified and acceptable under the circumstances. Nevertheless, even if the transgressors believe their acts to be justified, knowing that the victims (and possibly others in their moral community) perceive their behavior as immoral poses a threat to transgressors’ identity.

Research on the ‘Big Two’ in social psychology (see Abele et al., 2021) suggests that individuals’ and groups’ identities are perceived and judged along two fundamental dimensions: a ‘vertical’ dimension, representing traits such as agency and competence that are crucial for ‘getting ahead,’ and a ‘horizontal’ dimension, representing traits such as sociability and morality that are crucial for ‘getting along.’ According to the needs-based model of reconciliation (Shnabel & Nadler, 2008) transgressors experience threat to the ‘horizontal’ dimension of their identity (whereas their victims experience threat to the ‘vertical’ dimension). Because people generally wish to maintain their positive identity (Steele, 1988), and because those who violate the moral standards of their community face the risk of social rejection (Tavuchis, 1991), transgressors experience a heightened need to restore their positive moral identity.

How can this need be addressed? One way through which transgressors can restore their moral identity is apologizing to the victims and receiving their forgiveness. Theorizing on forgiveness has likened it to a ‘gift’ that victims grant to those who have offended them (Enright et al., 1998), which mitigates the moral inferiority engendered by the role of transgressor and reassures that the transgressors belong to the moral community to which their membership was questioned (Exline & Baumeister, 2000). The problem with apologies, however, is that they are risky. This is because, rather than reciprocating with forgiveness the victim might use the apology, which serves as an acknowledgment of the transgressor’s ‘moral debt’ to the victim (Minow, 1998), to further reproach the transgressor. Therefore, despite the common view of apologizing as a moral imperative (Benziman, 2009) and evidence for its effectiveness in promoting reconciliation (Schumann, 2018), transgressors often avoid taking this risk. They choose, instead, to defend their moral identity through moral disengagement, such as denying their culpability or minimizing the harm caused by the transgression (Bandura, 1999).

Optimistically, however, the needs-based model of reconciliation puts forward the hypothesis that addressing transgressors’ need to restore their moral identity may increase their readiness to take the risk involved in apologizing and increase their reconciliation efforts. In the next section I will review the research that supports of this hypothesis.

 **Addressing transgressors’ need for positive moral identity can open them to apology and reconciliation**

The hypothesis that satisfying transgressors’ need for moral acceptance can open them to reconciliation was first put into an empirical test in a series of experiments using diverse methods, including transgressions ‘orchestrated’ in the lab, recollection of real-life transgressions, and role-playing scenarios (see Shnabel & Nadler, 2008). For the present chapter’s purposes, I will briefly present the findings of the large-scale registered replication of one study in this series (Baranski et al., 2020), which was conducted as part of the Many Labs 5 project (Ebersole et al., 2020).

Participants in this replication study (*N* = 2,738) were undergraduates of seven American universities and one European university. They read a vignette about a recently unemployed college student who, upon returning from a 2-week family visit, learns that their roommate found a new roommate who could commit to paying the next year’s rent, and therefore the college student had to move out. Participants were randomly assigned either to the victim condition, in which they imagined themselves in the shoes of the roommate who had to leave the apartment, or to the transgressor condition, in which they imagined themselves in shoes of the roommate who stayed in the apartment.

Note that we intentionally chose a vignette that does not reflect vicious or inexcusable behavior: The transgressors can justify their behavior by claiming that staying with a roommate who cannot commit to pay the rent might get them in trouble. This choice stemmed from our theoretical stance that purely ‘evil,’ unjustifiable behavior (such as having sex with a friend’s fiancé, as in a vignette used in previous research by Gonzales et al., 1992) is rare, and our wish to use a vignette that simulates real-life transgressions. Indeed, this particular vignette was developed based on a pilot study in which U.S. undergraduates wrote about a transgression experienced in their own lives.

After the assignment to social roles (victims or transgressors), participants completed self-reported measures of their sense of agency, moral image, need for empowerment (wish to have greater control over the situation), need for moral acceptance (e.g., wish that the other roommate would perceive them as a moral person), and willingness to reconcile with their roommate. Next, participants received the second part of the vignette. It described a class on interpersonal dynamics, taken one week after the conflict, which both roommates attended and in which participants provided each other with feedback about their intellectual competencies and interpersonal skills. In the ‘empowerment’ condition the roommate was said to give the participant highly positive feedback about their intellectual competencies, whereas in the ‘moral acceptance’ condition the roommate was said to give the participant highly positive feedback about their interpersonal skills (e.g., warmth and niceness). Then, participants completed once again the measures of their sense of agency, moral image, and willingness to reconcile with their roommate.

The results for participants assigned to the transgressor condition[[1]](#footnote-1) revealed that in the first (‘before’) measurement they reported a lower moral image and a higher need for moral acceptance as compared to participants in the victim condition. Participants in the transgressor condition also reported a higher willingness to reconcile than participants in the victim condition, consistent with Baumeister’s (1997) observation that transgressors find it easier to ‘move on’ than victims. Comparing the ‘before’ and ‘after’ measurements revealed that the moral acceptance (but not the empowerment) message improved transgressors’ moral image. Moreover, the change in transgressors’ willingness to reconcile was higher in the moral acceptance than in the empowerment condition, such that ultimately transgressors’ willingness to reconcile was higher following the receipt of a morally accepting as compared to an empowering message from their victims. These findings suggest that restoring transgressors’ moral identity through an appropriate message from the victim can increase their goodwill towards the victim — even in comparison to a message that is highly positive in tone yet does not directly refer to their moral identity.

Further empirical support for the positive effect of restoring transgressors’ moral identity on their goodwill towards their victims was provided by research conducted within the framework of self-affirmation theory (Steele, 1988). According to this theory, behavioral or cognitive events that bolster the perceived integrity of the self (i.e., the person’s overall image as adequate) can protect individuals from psychological threats encountered in their environment. Self-affirmation interventions, which commonly involve short writing exercises (typically instructing participants to write about their most important value; McQueen & Klein, 2006), have been consistently found to reduce individuals’ defensive responses to psychological threats (for a review see Cohen & Sherman, 2014). For example, writing about their important values reduced smokers’ defensiveness in response to threatening health-related information (Crocker et al., 2008).

Applying this logic to the context of interpersonal transgressions, Schumann (2014) hypothesized that transgressors tend to offer their victim defensive, unsatisfying apologies because an apology inherently associates the transgressors with their wrongful behavior, thus further endangering their already shaken sense of being a good person. If so, then self-affirmation exercises may reassure the transgressors that they are ‘good people’ and improve the quality of the apologies they offer to their victim. To test this hypothesis, Schumann (2014) instructed participants to think about something that they had done that offended or hurt somebody else, and then write down what they would say to that person had he or she been there right now. As expected, affirmed (vs. non-affirmed) participants wrote messages that included fewer defensive strategies (excuses, justifications, victim blaming, and minimization of the harm) and more genuinely apologetic elements (expressions of remorse, responsibility taking, offer of repair, promise of forbearance, and request for forgiveness). The effect persisted even when controlling for mood, thus allowing to rule out positive mood as an alternative explanation.

A remaining question, however, was whether the positive effect of self-affirmation on transgressors’ conciliatory messages indeed stemmed from the restoration of their moral identity, or whether it was driven by their enhanced sense of connectedness due to the reassurance of their social bonds. The latter possibility is consistent with findings that when completing self-affirmation exercises participants typically write about their sense of belonging to significant others (such as friends and family members, Shnabel et al., 2013) and that sense of belonging may serve as a ‘symbolic shield’ that helps people cope with psychological threats (Mikulincer & Shaver, 2007). Indeed, both morality and belonging can be viewed as components of the ‘horizontal dimension’ of the ‘Big Two,’ representing one’s communion and getting along with others (Abele et al., 2021). Nevertheless, the restoration of one’s moral identity vs. one’s sense of belonging and connectedness to significant others can be viewed as representing two distinct avenues for encouraging reconciliation.

To distinguish between these two avenues, Woodyatt and Wenzel (2014) compared between different types of self-affirmation exercises. Participants in Woodyatt and Wenzel’s research had committed an interpersonal transgression a couple of days prior to study participation. They were then assigned to different affirmation conditions: In the morality affirmation condition, participants affirmed the value violated by the transgression by explaining why they felt this value was important to them and describing a time in the past in which they had behaved consistently with this value. In the two other affirmation conditions, participants affirmed either their sense of belonging (by writing about a time in which they felt loved or accepted), or a value that was important to them yet unrelated to the transgression. There was also control, no-affirmation condition in which participants wrote about the activities they were going to complete that day. The results revealed that, compared to the other experimental conditions, affirming the value violated by transgression led participants to process their feelings of shame, rather than leaving it unresolved. Processing their shame led, in turn, to genuine self-forgiveness; namely, transgressors’ effortful act of processing their wrongdoing, as compared to pseudo, defensive self-forgiveness characterized by simple lack of self-condemnation. Genuine self-forgiveness, in turn, increased participants’ trust that they would behave better in the future as well as their readiness to reconcile with their victim one week following the affirmation.

Notably, these results diverge from previous findings on self-affirmation interventions, in which threats to one’s identity and self-worth in one domain were effectively removed through the affirmation of one’s identity and self-worth in other domains (Cohen & Sherman, 2014). For example, Black students’ affirmation of their sense of belonging (representing the ‘getting along’ dimension of their identity) buffered them against the impairment to academic performance (representing the ‘getting ahead’ dimension of their identity) resulting from stereotype threat (the fear of confirming the negative stereotype about their group’s intelligence) (Shnabel et al., 2013). Woodyatt and Wenzel (2014) explain this discrepancy by arguing that in the traditional self-affirmation research, self-affirmation exercises are used to encourage perseverance in the face of negative feedback (e.g., about one’s academic performance). Yet “moral failure is not about poor performance at a valued task or activity” (p. 132), because one’s moral identity is intimately linked to their feeling as worthy and adequate (moreso than performance in academic tasks, for example). This argument underlines that morality is the most important dimension in people’s identity (see Leach et al., 2007), which makes it unique and non-fungible in nature.

Taken together, Baranski et al.’s (2020), Schumann’s (2014) and Woodyatt and Wenzel’s (2014) findings are all consistent with the notion that ‘what goes on between people, cannot be separated from what is going on within people’ (Gopin, 2004): Transgressors need to feel good about themselves, that is, perceive themselves as good and moral people in order to be kind and moral towards their victims.

**Unresolved issues concerning the positive effect of restoring transgressors’ moral identity**

The research reviewed so far has demonstrated that the restoration of transgressors’ moral identity either through morally accepting messages from the victim, or through self-affirmation exercises (especially if focused on the values breached by the transgression), increases transgressors’ genuinely conciliatory tendencies towards their victims. Research on identity restoration through messages conveyed by third parties, however, revealed a more complex picture. In this research (Shnabel et al., 2014, Study 2), participants read a vignette about an employee who took a 2-week leave from work due to maternity leave or military reserve duty.[[2]](#footnote-2) Upon returning to the office, the employee learned that a colleague who temporarily filled their position was ultimately promoted to their job, whereas they themselves were demoted. Participants assigned to the victim role were asked to imagine themselves as the demoted employee, and participants assigned to the transgressor role—as the promoted employee.

After the assignment to roles, participants were randomly assigned to three experimental conditions. In the control/no-message condition, the vignette ended at this point. In the two ‘message source’ conditions, the vignette continued to describe a subsequent staff meeting in which employees gave each other feedback. Transgressors assigned to the ‘other party’ condition read that the victim praised their interpersonal skills (e.g., mentioned incidents in which they were considerate of others at the workplace). For transgressors in the ‘third party’ condition, the same message was said to be conveyed by a colleague who was not involved in the conflict.[[3]](#footnote-3) Finally, participants completed a set of measures including moral image in the eyes of the victim, moral image in the eyes of the third party (the colleague not involved in the conflict), trust in the victim, and willingness for reconciliation. While a detailed review of the results is beyond the scope of the present chapter, Figure 1 presents the findings most relevant to the current discussion:

**+**

**+**

**+**

Morally accepting message from the victim

**+**

Moral image in the eyes of the victim

Trust in the victim

Willingness to reconcile

**+**

Morally accepting message from a third party

**\_**

Moral image in the eyes of the third party

Willingness to reconcile

*Figure 1*. Main findings of mediation analysis: A morally accepting message from the victim restored transgressors’ moral image in the eyes of the victim and trust in the victim, which was associated with the transgressor being *more* willing to reconcile (upper part of the figure). A morally accepting message from an uninvolved third party restored transgressors’ moral image in the eyes of the third party, which was associated with the transgressor being *less* willing to reconcile (lower part of the figure).

The results revealed that a morally accepting message from the victim restored transgressors’ trust in the victim (i.e., belief in the victim’s good intentions) and moral image in the victim’s eyes (but not in the eyes of the third party). Both trust in the victim and moral image in the eyes of the victim were associated with transgressors’ greater readiness for reconciliation. In contrast, a morally accepting message from a third party failed to affect transgressors’ trust in the victim or moral image in the victim’s eyes. It did restore transgressors’ moral image in the eyes of the third party, yet such restoration was *negatively* associated with transgressors’ readiness for reconciliation. This result points to the potentially adverse effect on reconciliation of morality restoration interventions by third parties. From a broader perspective, it suggests that there are boundary conditions to the positive effect on reconciliation of affirming and restoring transgressors’ moral identity.

**Future research directions**

An intriguing and important avenue for future research is to identify the conditions under which morality restoration effectively promotes transgressors’ conciliatory tendencies, as well as the conditions under which it might hinder them. This may be achieved by integrating the assumptions of the needs-based model with the literature on moral licensing effects (e.g., Merritt et al., 2010), which shows that people’s belief that they are good and moral people sometimes frees them to behave selfishly or immorally. The logic of this literature can explain the abovementioned finding that a morally accepting message from a third party was associated with transgressors’ lower readiness to reconcile (Shnabel et al., 2014; Study 2). Possibly, the praise to their interpersonal skills by the third party provided the transgressors with ‘moral credits’ — by reassuring that their bad behavior toward the victim was balanced out by their other, positive behaviors, or with ‘moral credentials’ — by establishing their ‘accreditations’ as “good people” (see Monin & Miller, 2001).

A possible account for the discrepancy between the prosocial effects of morality affirmation reported in the reconciliation literature (except for Shnabel et al., 2014) and the antisocial effects reported moral licensing literature is that the effect of moral affirmation depends on the saliency of a threat to transgressors’ moral identity. Perhaps transgressors show more conciliatory tendencies when *both* the threat to *and* the affirmation of their moral identity become salient at the same time. This could be the case, for example, when the victims themselves convey the morally reassuring message (as found by Baranski et al., 2020), or when transgressors write about why, despite the fact that they violated it, they do adhere to the value breached by the transgression (as found by Woodyatt & Wenzel, 2014). In contrast, when the affirmation occurs while the threat to transgressors’ identity is not salient, for example, when the morally reassuring message is conveyed by a third-party who was not involved in the conflict (as found by Shnabel et al., 2014), it might lead to more defensive, non-conciliatory responses.

Indeed, in the studies reported in the moral licensing literature, the moral affirmation is typically not directly linked to the immoral or unethical behavior. To illustrate, participants may believe that they complete a series of unrelated experiments, while in fact the first ‘experiment’ constitutes the experimental manipulation (e.g., a writing task through which they affirm their moral identity) and the last ‘experiment’ constitutes the measurement of (im)moral behavior (e.g., defection in a commons dilemma; Sachdeva et al., 2009). While theoretically plausible, the possibility that the effect of moral affirmation depends on the saliency of a moral threat awaits direct empirical verification.

Notably, however, one of the difficulties that hinders the empirical testing of this theoretical account (or, in fact, other hypotheses about transgressors’ moral behavior) is that assigning participants to the role of transgressors in lab experiments, which is the most preferred method in social psychological research, is tricky. While the victim’s role is passive and therefore easy to produce in the lab (e.g., by having participants being mistreated by a confederate), the transgressor’s role is active and therefore harder to orchestrate—as it paradoxically requires leading participants to transgress against others, yet making them do so *on their own volition*.

Classic paradigms that successfully lead most participants to transgress against other people, such as the “teacher-student” paradigm (Milgram, 1974), are ethically questionable. Recall of real-life transgressions has the advantage of participants’ high degree of psychological involvement, yet the conclusions might be limited to past conflicts that have already been thought about and processed, and standardization is compromised because the transgressions recalled by participants vary in terms of severity, foreseeability, and elapsed time since the event. Vignette studies allow control and standardization, but compromise external validity because participants respond to imagined scenarios rather than to real-life events, and the dependent variables are assessed through self-reports rather than based on participants’ actual behavior. Avoiding reliance on self-reports is possible through using behavioral economic games; however, this often necessitates focusing only on those participants whose behavior was transgressive (e.g., participants who breached their partners’ trust in a modified trust game; Leunissen et al., 2012), which might result in a selection bias.

Encouragingly, technological advances allow the implementation of new methods, such as computerized pseudo-dyadic ‘interactions,’ in which participants falsely believe that they are interacting with a partner, that are programmed in a dynamic way. In one study, for example (SimanTov-Nachlieli & Shnabel, 2014; Study 1), transgressors were ‘produced’ in the lab by asking participants to allocate valuable resources between themselves and an (ostensive) partner and then giving them feedback on their allocation behavior. The feedback was pre-programmed such that participants who allocated the resources unequally were informed that the norm was to allocate the resources equally, whereas participants who allocated the resources equally were informed that the norm was to allocate the resources “generously.” Thus, regardless of their allocation behavior, participants perceived themselves as violating common moral standards. This paradigm, which allows to study participants’ actual behavior under standardized conditions without the risk of selection bias, was originally used in the lab. Yet there is evidence that pseudo-dyadic ‘interactions’ can be effectively utilized in large-scale online platforms, such as MTurk (Summerville & Chartier, 2013). Hopefully, these technological tools and platforms will make it less challenging to experimentally study transgressors’ moral behavior in the future.

**Practical implications**

Identifying the circumstances under which morality restoration strategies by different sources, that is, the transgressors themselves, their victims, or third parties, can promote (or hinder) the transgressors’ efforts to apologize, compensate and reconcile with their victims has practical implications for mediators and facilitators of restorative justice practices. Such practices, which focus on rectifying the relationships and personal connections damaged by the transgression, rather than merely punishing the transgressors (Wachtel & McCold, 2001) are increasingly used in the legal and educational systems, as well as among families and communities (Boyes-Watson, 2008).

Participation in restorative justice procedures is typically not mandatory. For example, perpetrators of criminal offences in the Netherland voluntary choose whether to participate in victim-offender mediation (Jonas-van Dijk et al., 2020). The findings reviewed in this chapter suggest that transgressors’ motivation to take part in these procedures should be higher if they provide them with an opportunity to address their need for moral acceptance. A study among 91 victim-offender mediation cases from a Dutch mediation agency (Zebel et al., 2019) provided initial support for this possibility. This study revealed that the need to restore their moral identity was an important underlying factor in offenders’ decision to participate in VOM (victim-offender mediation) and intention to apologize and help their victims. Future research may examine whether the success of VOM (e.g., in reducing the risk of reoffending) is enhanced if encounters are structured such that offenders have an opportunity to address their need to restore their moral identity. Future research may also examine whether some offenders; for example, those who are dispositionally high on perpetrator sensitivity (more prone to feel guilt when transgressing against others; Baumert & Schmitt, 2016), are more likely to benefit from receiving an opportunity for moral restoration. Such future research may provide valuable insights for practitioners who engage in restorative justice encounters and real-life reconciliation efforts.

**Conclusion**

Based on the assumptions of the needs-based model of reconciliation, this chapter has put forward three arguments. First, the experience of threat to their moral identity is a common reason for transgressors’ tendency to refrain from making sincere effort to reconcile and make amends to their victims. Second, removing the threat to transgressors’ moral identity, which can be achieved through self-affirmation exercises that restore their moral image or through morally accepting message from their victims, increases transgressors’ readiness to reconcile with their victims and offer them remorseful, non-defensive apologies. Third, that morally accepting messages from third parties were associated with transgressors’ lower willingness to reconcile suggests that there are boundary conditions to the positive effects on reconciliation of morality affirmation, which should be identified in future research. I hope that the present chapter, and volume, will incite additional research on this socially and practically important topic.

**References**

Abele, A. E., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2021). Navigating the social world: Toward an integrated framework for evaluating self, individuals, and groups. Psychological Review, 128, 290–314. [https://doi.org/10.1037/rev0000262](https://psycnet.apa.org/doi/10.1037/rev0000262)

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. Personality and Social Psychology Review, 3, 193–209. https://doi.org/10.1207/s15327957pspr0303\_3

Baranski, E., Baskin, E., Coary, S., Ebersole, C. R., Krueger, L. E., Lazarevic´, L. B., . . . Žeželj, I. (2020). Many Labs 5: Registered Replication of Shnabel and Nadler (2008), Study 4. *Advances in Methods and Practices in Psychological Science*, *3*, 405–417.

Baumeister, R. F. (1997). *Evil: Inside human violence and cruelty.* New York: Henry Holt.

Baumeister, R. F., Stillwell, A. M., & Wotman, S. R. (1990). Victim and perpetrator accounts of interpersonal conflict: Autobiographical narratives about anger. *Journal of Personality and Social Psychology*, *59*, 994-1005.

Baumert A., & Schmitt M. (2016) Justice sensitivity. In: Sabbagh C., & Schmitt M. (eds) *Handbook of social justice theory and research*. Springer. <https://doi.org/10.1007/978-1-4939-3216-0_9>

Benziman, Y. (2009). *Forgive and not forget: The ethics of forgiveness*. Jerusalem: Van Leer [in Hebrew].

Boyes-Watson, C. (2008). *Peacemaking circles and urban youth: Bringing justice home*. Living Justice Press.

Cohen, G.L., & Sherman, D.K. (2014). The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology*, *65*, 333–371.

Crocker, J., Niiya, Y., & Mischkowski, D. (2008). Why does writing about important values reduce defensiveness? Self-affirmation and the role of positive other-directed feelings. *Psychological Science, 19*, 740-747.

Dunbar, R.I.M. (2009). The social brain hypothesis and its implications for social evolution. *Annals of Human Biology*, 36, 562–572. doi: 10.1080/03014460902960289

Ebersole, C.R., Mathur, M.B., Baranski, E.,… Szecsi, P. (2020). Many Labs 5: Testing pre-data-collection peer review as an intervention to increase replicability. *Advances in Methods and Practices in Psychological Science, 3,* 309-331.

Enright, R. D., Freedman, S., & Rique, J. (1998). The psychology of interpersonal forgiveness. In R. D. Enright & J. North (Eds.), *Exploring forgiveness* (pp. 46-62). Madison, WI: University of Wisconsin Press.

Exline, J. J., & Baumeister, R. F. (2000). Expressing forgiveness and repentance: Benefits and barriers. In M. E. McCullough, K. I. Pargament, & C. E. Thoresen (Eds.), *Forgiveness: Theory, research and practice* (pp. 133-155). New York: Guilford Press.

Gonzales, M. H., Manning, D. J., & Haugen, J. A. (1992). Explaining our sins: Factors influencing offender accounts and anticipated victim responses. *Journal of Personality and Social Psychology, 62*, 958–971.

Gopin, M. (2004). *Healing the heart of conflict: 8 crucial steps to making peace with yourself and others.* Emmaus, PA: Rodale Books.

Jonas-van Dijk, J. , Zebel, S., Claessen, J., & Nelen, H. (2020). Victim–Offender Mediation and Reduced Reoffending: Gauging the Self-Selection Bias: Gauging the Self-Selection Bias. Crime and Delinquency, 66(6-7), 949-972. https://doi.org/10.1177/0011128719854348

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*, 234–249.

Leunissen, J. M., De Cremer, D., & Reinders Folmer, C. P. (2012). An instrumental perspective on apologizing in bargaining: The importance of forgiveness to apologize. Journal of Economic Psychology, 33, 215-222.

McCullough, M.E., Kurzban, R., & Tabak, B.A. (2013). Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences, 36*, 1-58.

McQueen, A., & Klein, W.M.P. (2006). Experimental manipulations of self-affirmation: A systematic review. *Self and Identity, 5*, 289-354.

Merritt, A. C., Effron, D. A., & Monin, B. (2010). Moral self-licensing: When being good frees us to be bad. Social and Personality Psychology Compass, 4, 344–357.

 https://doi.org/10.1111/j.1751-9004.2010.00263.x

Minow, M. (1998). *Between vengeance and forgiveness: Facing history after genocide and mass violence*. Boston: Beacon Press.

Mikulincer, M., & Shaver, P. R. (2007). Boosting attachment security to promote mental health, prosocial values, and inter-group tolerance. *Psychological Inquiry, 18*, 139-156. https://doi.org/10.1080/10478400701512646

Milgram, S. (1974). *Obedience to authority: An experimental view*. Harper Collins.

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. Journal of Personality and Social Psychology, 81(1), 33–43. https://doi.org/10.1037/0022-3514.81.1.33

Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning saints and saintly sinners: The paradox of moral self-regulation. Psychological Science, 20, 523–528.

Schönbach, P. (1990). *Account episodes: The management or escalation of conflict*. New York: Cambridge University Press.

Schumann, K. (2014). An affirmed self and a better apology: The effect of self-affirmation on transgressors’ responses to victims. *Journal of Experimental Social Psychology*, *54*, 89–96.

Schumann, K. (2018). The psychology of offering an apology: Understanding the barriers to apologizing and how to overcome them. Current Directions in Psychological Science, 27, 74–78. [https://doi.org/10.1177/0963721417741709](https://psycnet.apa.org/doi/10.1177/0963721417741709)

Shaver, K.G. (1985). *The attribution of blame: Causality, responsibility and blameworthiness*. New-York: Springer.

Shnabel, N., & Nadler, A. (2008). A needs-based model of reconciliation: Satisfying the differential emotional needs of victim and perpetrator as a key to promoting reconciliation. Journal of Personality and Social Psychology, 94, 116–132. [https://doi.org/10.1037/0022-3514.94.1.116](https://psycnet.apa.org/doi/10.1037/0022-3514.94.1.116)

Shnabel, N., Nadler, A., & Dovidio, J. F. (2014). Beyond need satisfaction: Empowering and accepting messages from third parties ineffectively restore trust and consequent reconciliation. European Journal of Social Psychology, 44, 126–140. [https://doi.org/10.1002/ejsp.2002](https://psycnet.apa.org/doi/10.1002/ejsp.2002)

Shnabel, N., Purdie-Vaughns, V., Cook, J.E., Garcia, J., & Cohen, G.L. (2013). Demystifying values-affirmation interventions: Writing about social-belonging is a key to buffering against stereotype threat. *Personality and Social Psychology Bulletin, 39,* 663-676*.*

Silk, J.B. (2002). The form and function of reconciliation in primates. *Annual Review of Anthropology, 31*, 21-44.

Steele, C.M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self *Advances in experimental social psychology* (Vol. 21: Social psychological studies of the self: Perspectives and programs, pp. 261-302). San Diego,CA: Academic Press.

Summerville, A., & Chartier, C. R. (2013). Pseudo-dyadic “interaction” on Amazon’s Mechanical Turk. *Behavioral Research Methods, 45***,**116–124. https://doi.org/10.3758/s13428-012-0250-9

Tavuchis, N. (1991). *Mea culpa: A sociology of apology and reconciliation*. Stanford, CA: Stanford University Press.

Wachtel, T., & McCold, P. (2001). Restorative justice in everyday life. In H. Strang & J. Braithwaite (Eds.), *Restorative justice in civil society*. Cambridge University Press.

Woodyatt, L., & Wenzel, M. (2014). A needs-based perspective on self-forgiveness: Addressing threat to moral identity as a means of encouraging interpersonal and intrapersonal restoration. Journal of Experimental Social Psychology, 50, 125–135. https://doi.org/10.1016/j.jesp.2013.09.012

Zebel, S., Kippers, S., & Ufkes, E. (2019). Herstel van het morele imago van daders als drijfveer voor bemiddeling: De ervaringen van bemiddelaars, *Tijdschrift voor Herstelrecht, 19*, 22-35. DOI: 10.5553/TvH/1568654X2019019004004 [in Dutch]

1. Readers interested in the results for participants in the victim condition are kindly referred to Baranski et al. (2020). [↑](#footnote-ref-1)
2. Both are common reasons for absenteeism, for women and men respectively, in the Israeli context in which the study was conducted. [↑](#footnote-ref-2)
3. Participants in the victim condition received empowering messages, which praised their professional skills, from either the transgressor or a third party. Readers interested in these results are kindly referred to Shnabel et al. (2014). [↑](#footnote-ref-3)