

## Research article

# Justice through consensus: Shared identity and the preference for a restorative notion of justice

MICHAEL WENZEL<sup>1\*</sup>, TYLER G. OKIMOTO<sup>2</sup>,  
NORMAN T. FEATHER<sup>3</sup> AND MICHAEL J. PLATOW<sup>4</sup>

<sup>1</sup>*Flinders University, Adelaide, Australia*

<sup>2</sup>*Yale University, New Haven, USA*

<sup>3</sup>*Flinders University, Adelaide, Australia*

<sup>4</sup>*Australian National University, Canberra, Australia*

### Abstract

*We propose a concept of restorative justice as a sense of justice deriving from consensus about, and the reaffirmation of, values violated by an offence (in contrast to punishment-based retributive justice). Victims should be more likely to seek restorative justice (and less likely retributive justice) when they perceive to share a relevant identity with the offender. In Study 1, when the relevant identity (university affiliation) shared with the offender was made salient (vs. not), participants found a consensus-based response more justice-restoring. In Study 2, when the group (company) shared with the offender was cohesive (vs. not), participants more strongly endorsed a restorative justice philosophy and, mediated by this, responded in consensus-restoring ways. In Study 3, when the offender was an ingroup (vs. outgroup) member, participants more strongly endorsed a restorative justice philosophy, fully mediated by sadness emotions. Copyright © 2009 John Wiley & Sons, Ltd.*

When somebody violates a rule or law, we often feel an urge to pay them back, to retaliate or punish. Indeed, our sense of justice may demand us to do so. This notion of justice is widely understood as just desert or retributive justice and, in a way, is also institutionalized in our Western criminal justice system. However, in recent decades an alternative paradigm has emerged in criminology and law, called restorative justice. There are various practices of restorative justice, but the prototypical form involves a getting together of victim, offender and perhaps community representatives, where the incident and its aftermath are worked through in interaction between the parties. Ideally, the parties come to an agreement about the harm done, the values violated and the compensation, restitution and/or punishment required.

We contend that retributive and restorative justice practices imply fundamentally different notions of justice. From a social psychological perspective, it would seem important to understand what determines whether people subscribe to a retributive or restorative justice notion. In terms of practical value, a more refined understanding of people's conception of justice should assist in fine-tuning justice practices in order to maximize a sense of fairness among the involved parties. This may help secure a higher level of legitimacy of our regulatory institutions and authorities—both in the legal arena and in organizational contexts, and regardless of whether the transgressions are of an interpersonal nature or occur in the context of large-scale intergroup conflicts. Our present empirical contribution focuses on transgressions in the interpersonal domain, where offenders and victims are considered as individuals or individual members of a group.

\*Correspondence to: Michael Wenzel, School of Psychology, Flinders University, GPO Box 2100, Adelaide SA 5001, Australia.  
E-mail: michael.wenzel@flinders.edu.au

We examine two psychological conceptions of justice, retributive and restorative justice, and examine common identity as a key variable affecting the preference for each.

## RETRIBUTIVE AND RESTORATIVE JUSTICE

A transgression is understood as the violation of formal rules and laws or informal expectations and accepted standards of behaviour, in the course of which another party is victimized, harmed and deprived of entitlements. As Mikula (2003) argues, an injustice is likely to be perceived when an offender is considered responsible for the violation of another's entitlement that cannot be justified through other, higher-order ethical motives. Most, if not all, justice theories assume that experiences of injustice generally elicit the motivation to restore justice (e.g. Lerner, 1980; Walster, Walster, & Berscheid, 1978).

Indeed, research by Carlsmith (2006; Carlsmith, Darley, & Robinson, 2002; Darley, Carlsmith, & Robinson, 2000) suggests that the predominant motive following a transgression is the restoration of justice. However, these authors solely focus on punishment as a means to restoring justice. In line with a notion of just desert, Carlsmith and colleagues argue that people feel the urge to punish offenders because they believe offenders *deserve* to be punished in proportion to their wrongdoing. Punishment means the infliction of physical, material, psychological or social harm on the offender, which matches the negative moral quality of the offence and thus restores a moral balance (Carlsmith et al., 2002; Feather, 1999).

According to this notion of retributive justice, first, *penalizing the offender* is considered necessary, and often sufficient, to restore justice (in the strong version of 'positive retributivism'; see Duff, 2001). It is assumed punishment restores the moral balance that was upset by an offence. Second, the punishment has typically a degrading quality that lowers the offender's status or reputation, or *stigmatizes the offender*. From this perspective, a punishment that gives offenders the role and image of a good-hearted person as it were (e.g. community service) is misplaced, as it is inconsistent with a punishment's function of lowering the status of offenders and expressing censure (Kahan, 1996). Third, the punishment *per se* restores the justice; the punishment does not have to elicit feelings of remorse or contrition in the offender in order to be considered effective in restoring justice. Although there is evidence that jurors and victims impose more lenient punishments on remorseful offenders (Robinson, Smith-Lovin, & Tsoudis, 1994; Ohbuchi, Agarie, & Kameda, 1989), from a retributive justice perspective a more severe punishment is perfectly capable of restoring justice when offenders do not show any contrition. Nor do offenders have to agree to the punishment; it is sufficient (perhaps even necessary) that the punishment is *imposed* on the offender (see Duff, 2001). In short, retributive justice is *justice restoration through the imposition of punishment on the offender*.

Restorative justice has emerged in recent decades as a new paradigm in criminology and law (e.g. Bazemore, 1998; Braithwaite, 1998), but has attracted only little attention in social psychology so far (e.g. Cohen, 2001; Gromet & Darley, 2006; Morrison & Ahmed, 2006; Okimoto & Wenzel, 2008; Wenzel, Okimoto, Feather, & Platow, 2008). Nonetheless, related constructs of more constructive responses to rule-breaking have seen increased interest in social and organizational psychology lately, specifically around the concept of forgiveness (e.g. Aquino, Tripp, & Bies, 2006; Exline, Worthington, Hill, & McCullough, 2003; Worthington, 2005). Although restorative justice practices may involve retributive elements (and in practice the two justice notions may overlap; Daly, 2000), we posit that restorative justice is quite a different philosophy of restoring justice after transgressions. Specifically, it understands an offence as a conflict between offender, victim and the wider community, which therefore needs to be resolved in interaction between these parties (Christie, 1977; Marshall, 1998). Ideally, the parties develop a shared understanding of the wrongdoing, the victim's hurt, the offender's responsibility and the reparation or punishment required.

Restorative justice, first, commonly gives both victim and offender (and potentially representatives of their community) a *voice* in the process, for them to express their views and emotions. Second, based on a proper understanding of what the extent of the harm is, the process is designed to make the offender accept *accountability* for their actions and express a sincere *apology* to the victim. Third, the victim is in turn (at least implicitly) encouraged to express a *willingness to forgive* the offender and show basic *respect* to the offender as a human being generally capable of redemption and moral transformation (see Govier, 2002). Fourth, both parties are encouraged to decide together, *bilaterally*, on what is a suitable punishment for the offender and/or a suitable compensation for the victim. Note that punishment may be part of restorative justice but is not a necessary element; the parties may decide that punishment is required to underline their consensual disapproval of the offender's actions. In short, essential to restorative justice is a shared understanding of the harm the

offence has done and the values it violated and, through this consensus, to reaffirm those values: This is *justice restoration through consensus and the reaffirmation of values*.

### **CONCERNS ABOUT STATUS/POWER AND VALUES**

Which processes determine whether a retributive or restorative justice notion is adopted following a transgression in a given context? The answer partly lies, we contend, in the meaning that an offence has for those who are, directly or indirectly, affected by it (Wenzel et al., 2008). Specifically, the literature offers two possible accounts as to why people want to punish offenders (see Vidmar, 2000). First, the offender, through the offence, is perceived to have taken advantage of victim and society, expressed contempt for them and taken status and power away from them (Heider, 1958). The essential meaning of the offence is here the offender's usurpation of power and status, going hand in hand with feelings of humiliation and anger on the side of victim and community. In order to restore justice, the status/power relation needs to be re-established. The offender has to be 'cut down to size', disempowered and humiliated. According to this account, punishment fulfils exactly this function: Disempowerment and status reduction of the offender (Bies & Tripp, 1996; Vidmar, 2000). Second, an offence is the violation of a value consensus and disregard for a collectively shared view of what is right and wrong (Durkheim, 1964). Here, the essential meaning of the offence is the violation of shared values, which is likely to go hand in hand with feelings of disappointment or sadness in those affected by the transgression. In order to restore justice, the social consensus needs to be re-established and the values reaffirmed. It is assumed punishment does exactly that: It communicates social disapproval of the offender's actions and reaffirms a value consensus (Okimoto & Wenzel, 2009; Vidmar, 2000).

While both accounts have been made for the function of punishment, we argue that the status/power account is actually the more distinct motivation underlying the retributive justice notion, whereas the value reaffirmation concern is the more powerful motivation underlying the restorative justice notion (Wenzel et al., 2008; Wenzel & Thielmann, 2006). The imposition of punishment, as the defining element of retributive justice, is likely to reduce the status and power of the offender; it is therefore well suited to undo the status/power imbalance caused by the transgression. Likewise, concomitant to the frustration of status and power concerns are anger-related feelings (see Lazarus, 1991), which have been shown to be related to punitive and vengeful reactions (Averill, 1982). In contrast, the seeking of consensus, as a defining element of restorative justice, is instrumental to the reaffirmation of shared values and is therefore a particularly suitable response to value concerns. Moreover, disappointment about the offender's failure to conform to presumably shared values, and sadness about the harm to the group, are likely to elicit restorative responses that attempt to reintegrate the offender and heal the group (Okimoto, Wenzel, & Feather, 2009).

To summarize, different from a retributive justice notion, which is likely to be based on concerns about status and power, and feelings of anger and resentment, we assume that the endorsement of restorative justice in response to a transgression is motivated by concerns about group values and feelings of disappointment and sadness.

### **THE ROLE OF GROUP IDENTITY**

These considerations about the importance of the meaning of an offence lead us to specific predictions about the contextual conditions under which the re-establishment of a consensus is important and a restorative justice notion more likely to be endorsed. We argue this depends on the victim's perceived social identity in relation to the offender (Wenzel et al., 2008; Wenzel & Thielmann, 2006).

Previous research has suggested that the perceived ingroup or outgroup membership of the offender affects the level of punitivity in responses to transgressions. However, two conflicting propositions have been made. On the one hand, it has been argued victims or jurors will be more lenient when the offender is a member of their ingroup, and more punitive to offenders regarded as members of a salient outgroup (Graham, Weiner, & Zucker, 1997; Kerr, Hymes, Anderson, & Weathers, 1995; Sommers & Ellsworth, 2000); this corresponds to a straightforward ingroup favouritism effect (see Brewer, 2007). On the other hand, it has been suggested that an offence may be regarded as worse when it is committed by an ingroup member and may attract more extreme condemnation and punishment by fellow ingroup members for the apparent rejection of ingroup norms (Vidmar, 2002); this corresponds to a black sheep effect, that is, the more negative

evaluation of ingroup deviants compared to others who are equally deviant but belong to an outgroup (Marques & Paez, 1994; see also Marques, Abrams, & Serôdio, 2001). While we do not attempt to resolve these opposing processes completely, our theoretical conceptualization provides a more differentiated view, namely in terms of the form of justice that is sought (Wenzel et al., 2008).

Specifically, because the perceived validity of values depends on consensus within a relevant social group (Haslam, McGarty, & Turner, 1996), concerns about the value implications of an offence should be particularly salient when the offender is part of one's salient ingroup. According to self-categorization theory, to the extent that self and other are categorized as members of the same group, we expect the other to agree with us on relevant beliefs and values (Turner, 1987). Only when we expect the other to agree on the same values does their transgression question the validity of those values. If there is no expectation to agree we might not expect any different from the offender, and the transgression is unlikely to question our values (Turner, 1987). Similarly, if there is no expectation to agree, the working out of a consensus with the offender (e.g. in the sense of restorative justice) would do little to validate one's values. Our general hypothesis therefore is that sharing a relevant identity with the offender will lead to stronger endorsement of restorative justice responses.

These predictions can be contrasted with processes more likely to underlie the endorsement of retributive justice. Because status and power issues between victim and offender are inherently comparative (the offender's gain in status or power is the victim's loss), the interpretation of an offence as a usurpation of status and power implies a relationship of competition or negative goal interdependence between the parties. Negative interdependence in turn presupposes that one distinguishes between self and other, as between self and non-self (Turner, 1981). Concerns about one's status and power might thus be more relevant if one conceptualizes oneself as different from the offender; and retributive justice might be more strongly endorsed when the parties have different social identities or lack a strong common identity.

## OVERVIEW OF STUDIES

In this research, we are interested in seeking evidence for a psychological conception of restorative justice, understood as justice through consensus (in contrast to retributive justice). In three studies we unpack our general hypothesis that a shared identity will lead to greater endorsement of restorative justice. In Study 1 we test whether the salience of a shared identity will make the establishment of consensus more critical for perceptions of justice. In Study 2 we test whether a shared identity will lead to stronger endorsement of a restorative justice philosophy, which in turn accounts for a stronger preference for consensus-restoring responses. In Study 3 we test whether a shared identity will lead to greater concerns about values and stronger feelings of sadness, which in turn account for a stronger endorsement of a restorative justice philosophy.

### STUDY 1

In the first study participants were given information about a purportedly real incident at their university. In a case of academic dishonesty, a student took advantage of another student and falsely took credit for what was to be shared work. The case was dealt with by the university, and participants were given faked records that documented the procedure and outcome of the case. Participants were asked to take the perspective of the victimized student and respond to questions about the case. Prior to the transgression case, in what was allegedly an unrelated study, participants indicated their level of identification with their university, and the salience of that identity was manipulated via a priming technique.

## METHOD

### Participants and Design

Participants consisted of 138 undergraduates at an Australian university (70% female), between 18 and 56 years of age, with an average age of 24 years. The study employed a 2 (identity salience)  $\times$  3 (justice process) between-subjects design.

Participants were randomly assigned to one of the six experimental conditions ( $n = 23$  per cell), and were under the expectation that the experimental session involved multiple independent research studies.

## **Procedure**

Participants were told that the first study in the session investigated feelings about their student experience at the university, and they were asked to complete a number of questions (including the identity salience manipulation). They were then told that the second, unrelated investigation was mandated and funded by the Student Services Division of the university to assess student reactions to university grievance policies. Students would read about an account of a grievance filed by a fellow student at the university, and would be asked about their opinions regarding the situation and how the university handled it.

Participants were given an 'incident report' completed by a student services staff member, detailing the incident. In the report, the staff member explained a possible violation of university plagiarism or academic honesty policy reported by a class lecturer. The lecturer had suspected problems between two students in his class who had been assigned to work together on a group project, and contacted the students to investigate. One of the students (the victim) returned the lecturer's email, reporting that her group project partner (the offender) had lied to her about a family illness and tricked her into completing both portions of the assignment. Upon questioning the offender, she admitted to the class lecturer that she had deceived her assignment partner. The incident report stated that the lecturer had referred the conflict to Student Services. The victim's email to the lecturer detailing the situation was also attached to the incident report for participants to read.

After reviewing the incident details, participants read about the conflict resolution procedures implemented by Student Services. Apparently, the involved students were contacted and asked to attend a resolution meeting where the rules of the academic honesty policy were explained in detail. The description of the resolution also noted that, as punishment, the offending student would volunteer her time for student recruiting activities—a time consuming job ordinarily done by work-study students for pay. After reading all parts of the incident report, participants completed a number of questions assessing their opinions on the situation and how it was handled by Student Services.

## **Independent Variables**

### *Identity Salience*

The salience of participants' university identity was manipulated by priming their similarities (vs. differences) from other students of their university (Haslam, Oakes, Reynolds, & Turner, 1999; Wenzel, 2002). In the low salience condition, participants were instructed to 'think of what makes you different from other students at [the university]—what makes you an individual?' In addition, they were asked to complete the following sentence in three different ways: 'I am different from other [university] students because I am...'. In the high salience condition, on the other hand, students were instructed to 'think of what you share in common with other students at [the university]—what similarities make you a group?' In addition, they were asked to complete the following sentence in three different ways: 'I am similar to other [university] students because we all...'.

### *Justice Process*

The justice processes used in the resolution were manipulated by altering the resolution description in the Student Services incident report. In all conditions, the punishment (unpaid service to the university) was identical; however, the procedures used to come to that decision differed between conditions. In the *no consensus attempt* condition, 'after carefully weighing the implications of the offence', the student services staff member in charge of the resolution meeting 'made an objective decision regarding what I believed to be appropriate punishment.' It was also explained to the offender that she had no choice in the matter but to accept the punishment. In the *failed consensus* condition, the involved students 'expressed their emotions and views on the incident, but did not come to any agreement about appropriate student conduct or what they believed to be an appropriate punishment for the situation'. The students also 'did not seem to reach any shared understanding about the implications of the offence', and could not agree on a punishment. In the end, the Student Services

staff had to make a middle-ground punishment decision. Finally, in the *successful consensus* condition, the involved students again expressed their emotions and views on the incident, however, in this case they were able to come to 'an agreement about appropriate student conduct and what they believed to be an appropriate punishment for the situation'. The students 'seemed to reach a shared understanding about the implications of the offence', and agreed on the punishment.

## Dependent Variables

### *Manipulation Check*

Immediately after receiving the instructions for the identity salience manipulation (but prior to the justice process manipulation), participants were asked questions regarding their perceived similarity to other students at the university. These questions asked participants to rate their agreement (1 = *strongly disagree*, 7 = *strongly agree*) with three statements adopted from previous identity research (Jetten, Branscombe, Spears, & McKimmie, 2003; Jetten, Spears, & Manstead, 1997): 'I have a lot in common with other [university] students', 'I am a good example of a [university] student', and 'I am a typical [university] student'. Items were averaged for a composite manipulation check of identity salience ( $\alpha = .89$ ).

### *Measured University Identification*

Prior to the experimental manipulations, participants completed a measure of university identification, ostensibly as part of an unrelated study on the university experience. Participants were asked to rate their agreement (1 = *strongly disagree*, 7 = *strongly agree*) with five statements adopted from previous identity research (Wenzel & Thielmann, 2006): 'I identify with other [university] members', 'Being a member of the [university] community is important to me', 'When someone criticizes [the university], it feels like a personal insult', 'When I talk about [the university], I usually say "we" rather than "they"', and, 'I am a typical member of [the university]'. In addition, a 7-point Venn diagram measure assessed the degree to which participants perceived an overlap in their own identity and the university identity (Schubert & Otten, 2002; Tropp & Wright, 2001). The six items were averaged to form a single scale ( $\alpha = .89$ ).

### *Perceived Justice*

The primary dependent measures included assessments of satisfaction with the situation and its resolution, perceived fairness and perceived justice. Participants rated their agreement (1 = *strongly disagree*, 7 = *strongly agree*) with seven statements, including: 'I am pleased with the development of events', 'I am very dissatisfied with the situation' (reverse-coded), 'Fairness has been reestablished', 'My sense of fairness has been satisfied', 'The resolution has restored justice', 'The situation has been handled fairly', and 'The resolution to the situation is fair'. Items were adopted from previous research on justice resolutions (Okimoto et al., 2009). Factor analyses with varimax rotation indicated that all seven items loaded onto a single factor explaining 68% of the variance (eigenvalue = 4.74, factor loadings > .65); therefore, all items were averaged for a single measure of *perceived justice* ( $\alpha = .92$ ).

## RESULTS

### Manipulation Check

An initial *t*-test on the identity salience manipulation check indicated no significant difference between salience conditions,  $t(136) = 1.38$ ,  $p = .17$ . Additional analyses revealed that identity salience did have an effect on the manipulation check, but only when also considering measured identification (ID). A regression analysis using identity salience, measured ID and their interaction revealed a marginal main effect for identity salience,  $\beta = .10$ ,  $t(134) = 1.86$ ,  $p = .06$ , and a significant main effect for measured ID,  $\beta = .76$ ,  $t(134) = 13.83$ ,  $p < .001$ , in the expected directions. Moreover, there was a significant interaction between identity salience and measured ID,  $\beta = .16$ ,  $t(134) = 2.84$ ,  $p = .005$ .

Slope analyses indicated that for low identifiers (measured) the salience manipulation was not effective,  $\beta_x = -.06$ ,  $t = 0.48$ ,  $p = .48$ , but for high identifiers it was,  $\beta_x = .26$ ,  $t = 3.33$ ,  $p = .001$ . These results might not surprise given the relative weakness of the salience manipulation compared to individual differences in university identification, and the notion that salience of a self-category is an interactive function of situational fit (e.g. primed similarities vs. differences) and perceiver readiness (e.g. chronic identification; Oakes, 1987). Given this finding, after testing the initial hypothesized effects for the interaction between justice process and identity salience, we included measured identification using a regression approach.

### Analysis of Variance for Manipulated Variables

Analysis of variance of the dependent measure assessing perceived justice indicated no significant main effect for the identity salience manipulation,  $F(1,132) = 0.58$ ,  $p = .45$ . There was a significant main effect for the justice process manipulation,  $F(2,132) = 4.07$ ,  $p < .05$ ,  $\eta^2 = .06$ . This main effect was, however, qualified by a significant interaction between justice process and identity salience,  $F(2,132) = 5.12$ ,  $p < .01$ ,  $\eta^2 = .07$ . Analysis of simple effects indicated for participants in the low salience condition no differences between the no consensus ( $M = 4.35$ ,  $SD = 1.33$ ), failed consensus ( $M = 4.27$ ,  $SD = 1.00$ ) and successful consensus conditions ( $M = 4.23$ ,  $SD = 1.17$ ),  $F(2, 66) = .06$ , ns. In contrast, the justice process conditions differed significantly when university identity was made salient,  $F(2, 66) = 9.63$ ,  $p < .001$ . Participants rated the process involving successful consensus ( $M = 4.96$ ,  $SD = 1.11$ ) as significantly more just than the no consensus process ( $M = 3.79$ ,  $SD = 1.19$ ),  $t(44) = 3.45$ ,  $p = .001$  and significantly more just than the unsuccessful consensus process ( $M = 3.64$ ,  $SD = 1.05$ ),  $t(44) = 4.15$ ,  $p < .001$ . The no consensus and unsuccessful consensus conditions were not significantly different from one another,  $t(44) = -.46$ , ns. The results were in line with our prediction (see Figure 1).

### Considering Measured Identification

In order to analyse the effects of the independent variables in tandem (and interaction) with the continuous measure of university identification, we used multiple regression techniques outlined by Aiken and West (1991) for interactions between categorical and continuous variables. To facilitate the regression framework, predictor variables were centred and the three-level manipulated justice process variable was effect-coded. The effect-coded variable *successful consensus* indicates successful consensus between parties (given a value of .66) as compared to failed and no consensus conditions (−.33 each), while the variable *consensus attempt* indicates an attempt at consensus, whether eventually successful or not (.33 each), as compared to no attempt at consensus (−.66). Note that we used values that effectively centred the effect codes, so that the main effects could be interpreted as such even when an interaction term was included (Aiken & West, 1991). Statistics from the analysis are presented in Table 1.

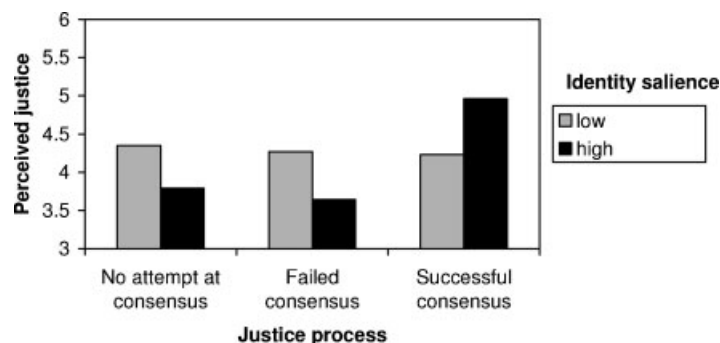


Figure 1. Measure of perceived justice: Interaction between manipulations of justice process and identity salience

Table 1. Regression results for dependent measure of perceived justice (Study 1)

Predictor	Unstandardized <i>B</i>	Standard error	Standardized $\beta$	<i>t</i> -value
(Constant)	4.16	.10	—	—
Measured identification (ID)	.04	.09	.04	.44
Identity salience	-.20	.20	-.09	-1.04
Consensus attempt	-.14	.24	-.06	-.59
Successful consensus	.72	.24	.28	3.06**
Identity salience $\times$ measured ID	.17	.19	.08	.92
Consensus attempt $\times$ identity salience	.15	.49	.03	.32
Consensus attempt $\times$ measured ID	-.70	.22	-.29	-3.13**
Successful Consensus $\times$ identity salience	1.44	.48	.28	3.03**
Successful consensus $\times$ measured ID	.60	.22	.24	2.67**

Note: Adjusted  $R^2 = .15$ ,  $F(9, 128) = 3.63$ ,  $p < .001$ . There were no significant three-way interactions, which were therefore omitted from the analysis. \*\* $p < .01$ .

The results showed a significant main effect of *successful consensus*,  $\beta = .28$ ,  $p = .003$ ; consensus resulted in generally more favourable justice perceptions. No other main effects were apparent. Consistent with the earlier ANOVA, a significant two-way interaction was found between *successful consensus* and *identity salience*,  $\beta = .28$ ,  $p = .003$ . Slope analysis indicated no significant effect of successful consensus when identity salience was low,  $\beta_x = .00$ , ns, whereas successful consensus resulted in more favourable justice perceptions when the identity was salient,  $\beta_x = .57$ ,  $p < .001$  (see Figure 2a). This effect was not further moderated by measured identification. The two three-way interaction terms (of

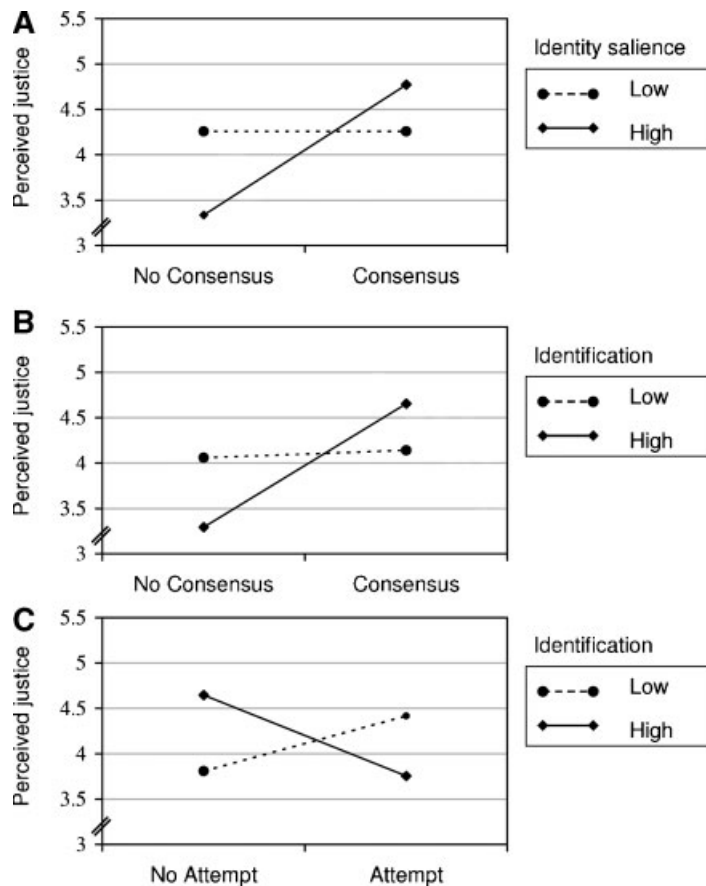


Figure 2. Slope analyses: Identity salience and measured identification as moderators of the effects of consensus-seeking on perceived justice

identity salience, measured ID and either consensus attempt or successful consensus) did not have a significant effect,  $t < 1$ , and were thus excluded from the analysis.

However, measured ID did show two significant two-way interaction effects. First, one significant interaction was found between *measured ID* and *successful consensus*,  $\beta = .24, p = .009$ . Slope analysis indicated exactly the same pattern as for identity salience, namely for low levels of measured ID successful consensus did not have a significant effect on perceived justice,  $\beta = .03$ , ns, but for high levels of measured ID successful consensus resulted in more favourable justice perceptions,  $\beta_x = .53, p < .001$  (see Figure 2b). A second interaction emerged between *measured ID* and *consensus attempt*,  $\beta = -.42, p = .002$ . While there was a marginally positive effect of consensus attempt for participants with low measured ID,  $\beta_x = .24, p = .071$ , the consensus attempt resulted in significantly *less* favourable justice perceptions for participants with high measured ID,  $\beta_x = -.35, p < .011$  (see Figure 2c).

## DISCUSSION

The study provides evidence that the establishment of consensus between the involved parties can contribute to the restoration of a sense of justice in victims of a transgression (Okimoto et al., 2009). This is consistent with our theoretical concept of restorative justice as the restoration of justice through consensus, which is distinct from the operation of retributive justice through the imposition of punishment on the offender. More importantly, the results show that consensus contributes to a sense of justice *in particular* among victims/participants whose shared identity with the offender (i.e. their shared university affiliation) had been situationally salient or chronically strong.

The similarity of the results for identity salience versus measured identification is striking. However, it is interesting that their moderation effects were independent from another, quite different from their interactive effect on the identity salience manipulation check. Situational fit and general importance of an identity seem to tap two different (non-redundant) aspects of identity (e.g. Ashmore, Deaux, & McLaughlin-Volpe, 2004); yet both operate in a similar fashion in qualifying the importance of consensus-based justice.

The additional finding that the mere attempt to achieve consensus showed the opposite interaction with measured identification was unexpected. We need to consider that this effect occurred while controlling for the success of the attempt. In effect, the result thus indicates that *failure* to achieve consensus led to a reduced sense of justice for participants who (*a priori*) identified strongly with their university. Again, this appears consistent with our theoretical analysis: Those who are chronically more strongly committed to the identity shared with the offender should be more strongly affected when consensus fails.

It needs to be highlighted that the attempt to reach consensus, which involved giving both parties an opportunity to express their views and having a say in the process, did not per se increase perceptions of justice. Hence, our findings do not merely reflect the operation of a fair process effect (Folger, Rosenfield, Grove, & Corkran, 1979), and our notion of restorative justice is not the same as procedural justice. Rather, it is crucial that the procedure leads to a consensus between the parties about the wrongdoing, the harm done, the values violated and options for their restoration.

## STUDY 2

The previous study showed that the re-establishment of consensus between the involved parties can restore a sense of justice, particularly among victims who, in the given context, define themselves in terms of an identity they share with the offender. We argue that this effect reflects the operation of restorative justice. In fact, we argue that consensus assumes a justice restoring function only to the extent that victims adopt a *restorative justice notion*. In this second study we tested this more explicitly by measuring restorative versus retributive justice as two different (context-specific) justice notions: Lay philosophies about how justice is to be restored after the wrongdoing. A restorative justice notion should be more strongly endorsed when victims perceive themselves as sharing a relevant identity with the offender; and mediated via the endorsement of a restorative notion, a shared identity should lead to the victim's support for responses to the injustice that are likely to facilitate consensus between the parties.

For this second study, we chose a different methodological approach, including a different manipulation of shared identity, to counter the possibility that our findings are method-specific artefacts and to demonstrate their more general validity. Specifically, we used a scenario methodology where participants were asked to imagine they were the owner of a delicatessen/catering company and they discovered that one of their employees was stealing from them. In an introductory paragraph, common identity was manipulated in terms of the cohesiveness of the situationally relevant group that included offender and victim, that is, the company. We assumed a cohesive group, being more meaningful and identity-relevant, would more likely elicit a sense of common identity with the offender (see Castano, Yzerbyt, & Bourguignon, 2003).

## METHOD

### Participants and Design

Participants were 83 university students (65% female) aged between 17 and 57 years of age, with an average age of 22 years. The design was a single factor between-subjects design manipulating cohesiveness (low vs. high). Note that slight variations in degrees of freedom in the statistical analyses are due to missing values.

### Procedure

Participants volunteered to participate in a questionnaire about rule breaking. They were instructed to imagine that they owned a small delicatessen and catering company, and to answer the questions in this study from that perspective. The questionnaire described an incident in which the participants, as the owner of the business, found out that one of their employees was stealing from their company and helping himself to food.

### Independent Variable

Cohesiveness was manipulated by a brief characterization of the company. In the low cohesiveness condition, employees were described as having low collegial spirit, all going their own way and rarely engaging with each other. In the high cohesiveness condition, the company was described as one where employees had high collegial spirit, saw themselves as one big family and engaged frequently with each other at a personal level.

### Dependent Variables

#### *Manipulation Check*

Immediately after the characterization of the company, but before the incident was described, two items measured the sense of common identity that was perceived to be prevalent in the company. First, a verbal measure asked, 'How strongly do you think the employees identify with your company?' (1 = *not at all*, 7 = *very strongly*). Second, a pictorial measure was used (based on Schubert & Otten, 2002): 'Which of the following pictures do you think represents best the relationship between your employees and your company?' There were four response options of different overlap between an X, representing any employee and a circle that represented the company. The two items were standardized and then averaged to form a measure of common identity ( $\alpha = .81$ ).

#### *Concrete Responses*

Participants were asked how they think they would react to the incident (1 = *unlikely*, 7 = *very likely*). Ten concrete responses were measured: Penalizing the offender (4 items; e.g. I withdraw my approval for the employee's request to

participate in a three-day training course the following week;  $\alpha = .78$ ); humiliating the offender (4 items; e.g. I shout abuse at the employee in front of his work colleagues;  $\alpha = .72$ ); unilateral imposition of punishment (2 items; e.g. The employee would have no choice but to accept the punishment that we impose on him;  $\alpha = .78$ ); victim voice (4 items; e.g. I and other staff explain to the employee how we feel about his behaviour;  $\alpha = .69$ ); offender voice (4 items; e.g. I let the employee explain the incident;  $\alpha = .91$ ); accountability (4 items; e.g. I ensure the employee takes responsibility for the harm done to the business;  $\alpha = .77$ ); apology (4 items; e.g. I demand an apology from the employee;  $\alpha = .79$ ); forgiveness (4 items; e.g. I offer my willingness to forgive the employee;  $\alpha = .65$ ); respect for offender (4 items; e.g. Despite what happened, I treat the employee with respect;  $\alpha = .82$ ); and bilateral decisions on punishment (1 item; We should find a penalty that all sides, including the employee and I, think is fair).

### Justice Notions

Based on previous research (Okimoto & Wenzel, 2008; Okimoto et al., 2009), levels of endorsement of a restorative ( $\alpha = .78$ ) and a retributive justice notion ( $\alpha = .88$ ) were measured by six items each (see Appendix). The two justice notions were practically uncorrelated ( $r = .03$ ). Exploratory factor analysis confirmed the distinction between the two justice notions, extracting two strong factors explaining 57% of the variance (a third factor with an eigenvalue greater than 1 was considerably weaker, and a Scree test provided evidence that a two-factor solution was more appropriate). All items loaded greater than .58 on their respective factor, with no substantial cross-loadings (see Appendix).

### Controls

As a measure of perceived *seriousness* of the offence, participants were asked to rate the incident on four attributes (1 = *not at all*, 7 = *very much*): Wrong, acceptable (reverse-scored), serious, no big deal (reverse-scored). The four items were averaged to obtain scale scores ( $\alpha = .72$ ). As a measure of perceived *intentionality*, the ratings on a further three attributes were averaged: Intentional, deliberate, wilful ( $\alpha = .83$ ). Finally, three items were included to measure the *justifiability* of the offender's actions (e.g. There might be good reasons for the offender's behaviour; 1 = *do not at all agree*, 7 = *strongly agree*;  $\alpha = .67$ ).

## RESULTS

### Manipulation Check and Controls

To check the effectiveness of the cohesiveness manipulation, the measure of perceived common identity was subjected to a *t*-test. As intended, in the high cohesiveness condition participants saw the employees more strongly as part of the company than in the low cohesiveness condition ( $M_s = .70$  vs.  $-.72$ ; scale scores were averaged *z*-scores),  $t(81) = 11.03$ ,  $p < .001$ ,  $d = 2.42$ . Further, it was important to establish that the identity manipulation was not confounded with a different appraisal of the incident in terms of its seriousness, intentionality or justifiability. Indeed, *t*-tests showed that participants in the high versus low cohesiveness conditions rated the incident as not differently serious ( $M_s = 5.99$  vs. 5.89),  $t(80) = .53$ , ns, intentional ( $M_s = 5.48$  vs. 5.61),  $t(78) = -.62$ , ns, or justifiable ( $M_s = 5.00$  vs. 4.81),  $t(81) = .52$ , ns. Generally, participants saw the offender's actions as clearly wrong and unacceptable.

### Identity Effects on Justice Philosophies and Concrete Responses

Across conditions, we first investigated the relationships between the two abstract justice notions and the more concrete behavioural options in response to the offence. The product-moment correlations are summarized in Table 2. As expected, retributive justice was positively related to penalizing the offender, humiliating the offender and unilateral punishment.

Table 2. Correlations between concrete responses to the transgression and notions of retributive and restorative justice (Study 2)

	1	2	3	4	5	6	7	8	9	10	11
(1) Retributive justice	—										
(2) Restorative justice	.03	—									
(3) Penalizing offender	.47***	-.04	—								
(4) Humiliating offender	.36***	-.17	.39***	—							
(5) Unilateral decision	.68***	.11	.55***	.21	—						
(6) Voice for victim	.34**	.32**	.37***	.28*	.34**	—					
(7) Accountability	.32**	.55***	.20	-.05	.44***	.46***	—				
(8) Apology	.04	.62***	.10	-.06	.16	.15	.59***	—			
(9) Forgiveness	-.36***	.41***	-.13	-.18	-.19	.11	.25*	.43***	—		
(10) Voice for offender	-.51***	.42***	-.23*	-.40***	-.30**	-.01	.24*	.39***	.66***	—	
(11) Respect for offender	-.32**	.36***	-.28*	-.40***	-.22*	.01	.20	.26*	.68***	.68***	—
(12) Bilateral decision	-.02	.42***	.04	-.10	.07	.19	.41***	.31**	.36***	.42***	.41***

\* $p < .05$ , \*\* $p < .001$ , \*\*\* $p < .001$ .

Unexpectedly, it was also positively related to giving voice to the victim and holding the offender accountable. Further, retributive justice was negatively related to some of the other, presumably consensus-oriented, responses: Offender voice, willingness to forgive, respect for the offender and bilateral decisions on punishment. In line with expectations, restorative justice was not related to the presumably retributive responses of penalizing and humiliating the offender and unilateral punishment. Rather, as expected, it was positively related to the remaining seven reactions that we expected to be more restorative. Overall, this provided good construct validation of our two justice notions.

Applying  $t$ -tests, we tested for the effects of the cohesiveness manipulation on the endorsement of retributive and restorative justice notions. As can be seen in Table 3, the restorative justice notion was, as predicted, significantly more strongly endorsed in the high than the low cohesiveness condition ( $M_s = 5.70$  vs. 4.86),  $t(80) = 4.07$ ,  $p < .001$ ,  $d = .90$ . Levels of endorsement of retributive justice did not differ significantly between the two experimental conditions, ( $M_s = 3.96$  vs. 3.49),  $t(64.02) = 1.53$ , ns (note,  $df$ s were adjusted for inequality of variances).

Similarly, we investigated whether the cohesiveness manipulation affected the likelihood that participants would engage in the different concrete responses. The results are summarized in Table 3. A series of  $t$ -tests showed no significant differences between cohesiveness conditions for the three presumably retributive responses, penalizing the offender, humiliating the offender and unilateral decision-making. However, among the presumably consensus-oriented responses, participants in the high (compared to low) cohesiveness condition were significantly more likely to demand accountability

Table 3. Means standard deviations, and  $t$ -statistics for endorsement of retributive and restorative justice notions and concrete responses to transgressions in groups with low versus high cohesiveness (Study 2)

	Low cohesiveness		High cohesiveness		$t$
	$M$	SD	$M$	SD	
Retributive justice	3.49	1.70	3.96	1.00	1.53
Restorative justice	4.86	1.11	5.70	.71	4.06***
Penalizing offender	3.00	1.62	2.90	1.15	-.33
Humiliating offender	1.53	.89	1.39	.50	-.90
Unilateral decision	4.20	1.93	4.37	1.60	.45
Voice for victim	3.24	1.41	3.43	1.40	.60
Accountability	4.75	1.34	5.87	.85	4.53***
Apology	4.52	1.43	5.99	.77	5.81***
Forgiveness	4.44	1.28	4.98	1.03	2.10*
Voice for offender	5.91	1.54	6.29	.83	1.41
Respect for offender	5.45	1.34	5.92	1.08	1.73
Bilateral decision	4.44	2.01	5.29	1.71	2.06*

\* $p < .05$ , \*\*\* $p < .001$ .

( $d = 1.00$ ), demand an apology ( $d = 1.28$ ), show willingness to forgive ( $d = .46$ ) and engage in bilateral decision-making ( $d = .46$ ); and, as a non-significant trend, they were also more likely to grant the offender voice ( $d = .31$ ) and respect ( $d = .39$ ).

### Mediation of Identity Effects

We investigated whether the effects of the cohesiveness manipulation on consensus-oriented responses were mediated by the endorsement of a restorative justice notion, as predicted in Hypothesis 2. Following Kenny, Kashy and Bolger's (1998) recommendations for examining mediation, a series of regression analyses showed (a) the cohesiveness manipulation of shared identity (as we have already seen) positively affected accountability ( $\beta = .45, p < .001$ ), apology ( $\beta = .54, p < .001$ ), forgiveness ( $\beta = .23, p = .037$ ) and bilateral decision-making ( $\beta = .23, p = .043$ ); (b) cohesiveness positively affected the endorsement of a restorative justice notion ( $\beta = .41, p < .001$ ) and (c) restorative justice, when controlling for the cohesiveness manipulation, was positively related to accountability ( $\beta = .44, p < .001$ ), apology ( $\beta = .48, p < .001$ ), forgiveness ( $\beta = .38, p < .001$ ) and bilateral decision-making ( $\beta = .39, p < .001$ ). When controlling for restorative justice, the effects of the cohesiveness manipulation on accountability ( $\beta = .27, p = .008$ ) and apology ( $\beta = .35, p < .001$ ) were reduced while still being significant; Sobel tests showed that the indirect effects on accountability ( $z = 3.00, p = .003$ ) and apology ( $z = 3.25, p = .002$ ) were significant, thus indicating partial mediation. Controlling for restorative justice, the effects of cohesiveness on forgiveness ( $\beta = .07, ns$ ) and bilateral decision-making ( $\beta = .06, ns$ ) were reduced to non-significance; Sobel tests showed that the indirect effects on forgiveness ( $z = 2.61, p = .011$ ) and bilateral decision-making ( $z = 2.64, p = .010$ ) were significant, indicating full mediation.

In addition, we performed a mediation analysis at a more aggregate level using the information on all responses. First, a second-order factor analysis (varimax rotation) on the 10 concrete responses yielded two factors with eigenvalues greater than 1, explaining 59% of the variance. The factor loadings (see Table 4) were consistent with a priori considerations in that accountability, apology, forgiveness, offender voice, respect and bilateral decision-making loaded substantially on one factor, which we may call consensus-oriented responses; and penalizing the offender, humiliating the offender and unilateral punishment loaded on the second factor, which we may call punitive responses. Unexpectedly, voice for victims loaded strongly on the punitive response factor and less so on the consensus-oriented response factor, and accountability loaded on both. This will be discussed later.

The factor scores for the consensus-oriented responses were then subjected to the same mediation analysis as before. Results showed that (a) the cohesiveness manipulation of shared identity increased the likelihood of consensus-oriented responses ( $\beta = .36, p < .001$ ); (b) as already seen, cohesiveness positively affected the endorsement of a restorative justice notion ( $\beta = .41, p < .001$ ) and (c) restorative justice, when controlling for cohesiveness, was positively related to consensus-oriented responses ( $\beta = .56, p < .001$ ). When controlling for restorative justice, the effect of cohesiveness on consensus-oriented responses was no longer significant ( $\beta = .15, ns$ ); a Sobel test showed that the indirect effect on consensus-oriented responses was significant ( $z = 3.36, p = .001$ ), indicating full mediation. By contrast, the factor scores for the punitive response factor were not significantly affected by the experimental manipulation ( $\beta = .17, ns$ ), and there

Table 4. Factor loadings for concrete responses to transgressions (Study 2)

	Factor 1	Factor 2
Penalizing offender	-.15	.74
Humiliating offender	-.37	.54
Unilateral decision	-.10	.78
Voice for victim	.19	.65
Accountability	.57	.63
Apology	.63	.35
Forgiveness	.81	-.14
Voice for offender	.82	-.29
Respect for offender	.78	-.31
Bilateral decision	.63	.19

was thus no effect to be mediated. The punitive response factor was strongly correlated with the retributive justice notion ( $r = .68, p < .001$ ), and more weakly correlated with the restorative justice notion ( $r = .26, p = .019$ ).

## DISCUSSION

Using a different methodology and operationalization of shared identity, the findings of this study are consistent with those of Study 1. When the common group of victim and offender was highly cohesive, and thus presumably salient and identity-defining, participants were more likely to engage in responses that should aid the re-establishment of a consensus between the affected parties. Specifically, they were more likely to seek an admission of accountability and an apology from the offender; and they were more likely to offer forgiveness and favoured a bilateral decision about an appropriate response. In addition, the present study provided evidence that the adoption of consensus-oriented responses in the case of a strong shared identity was mediated by the endorsement of a restorative notion of justice.

A restorative justice notion was conceptualized and measured as the view that the restoration of justice requires that the parties reaffirm the values violated by the transgression through renewed consensus. This was contrasted with a retributive notion according to which the restoration of justice requires that the offender be punished. The differential relationships that the two justice notions showed with consensus-oriented versus punitive responses (e.g. seeking an apology vs. penalizing the offender) provided a good construct validation of the two measures. Unexpectedly, though, victim voice was not only related to restorative justice but also to retributive justice. If we consider, however, that voice gives victims power (which they could use to humiliate the offenders or unilaterally impose a penalty on them), the result might not surprise. Likewise, accountability showed a positive relation to both a restorative and retributive justice notion. Holding someone accountable might have become a euphemism for making someone pay for their wrongdoing, which might explain the unexpected result.

In line with our prediction, we found that a restorative justice notion was more strongly endorsed when victim and offender shared membership in a cohesive group. However, the present study did not show that retributive justice was *less* strongly endorsed under high group cohesiveness, even though our introductory argument would suggest this. It might be the case that an experimental variation of the cohesiveness of the relevant group does not cover the whole spectrum from non-shared identity to shared identity. That is, in both low and high cohesiveness conditions, victim and offender were in fact part of the same group; only the 'groupiness' or entitativity (Campbell, 1958; Castano et al., 2003) of the group was varied. We reasoned that clearer effects, that is, a complete mirror image of effects for retributive and restorative justice, might be obtained with an explicit manipulation of the offender's inclusion versus non-inclusion in the victim's group. Such an approach would also overcome another ambiguity of the cohesiveness manipulation, namely that it might not have been social identity salience but rather good interpersonal relations and a concern for their maintenance that were responsible for the stronger endorsement of restorative justice.

## STUDY 3

In line with this reasoning, we manipulated shared identity in Study 3 differently from Study 2, while otherwise keeping the method the same in order to determine whether the type of identity manipulation would make a difference. That is, participants were again asked to imagine they were a shop-owner who found out that an employee stole from them. This time, however, the company was always portrayed as being cohesive (and identity-relevant), but the offending employee was either a regular employee of the company (an ingroup member) or only temporarily hired from another company (an outgroup member).

Furthermore, based on the evidence from Studies 1 and 2, we were satisfied that a restorative notion of justice is a valid concept that helps to understand what actions people prefer in order to restore justice. Our focus shifted therefore now to the processes, triggered by a shared (vs. non-shared) identity, which determine whether victims endorse a restorative or retributive justice notion. As may be recalled, we predicted that a shared identity would lead to stronger endorsement of restorative justice because of (mediated by) greater concern about values that were presumably shared with the offender, and mediated by greater sadness about the harm done to those values and one's group, identity and relationship with the

offender. In contrast, endorsement of retributive justice might be primarily affected by concerns about status and power as well as feelings of anger that are concomitant with the experience of humiliation and disempowerment.

## METHOD

### Participants and Design

Participants were 152 university students (59% female, with 3 cases without details on gender) aged between 17 and 64 years of age, with an average age of 23.6 years. The design was a single factor between-subjects design with the offender's group membership (outgroup vs. ingroup) as the experimental factor. The procedure was similar to Study 2.

### Independent Variables

Group membership of the offender was manipulated by providing information that the company had to take measures to deal with high seasonal demand. In the ingroup condition, it was said: 'In order to meet the demands of the holiday season, you reassigned one of your existing employees to catering in order to cover the company's needs—someone with experience, but who ordinarily worked only in food preparation'. In the outgroup condition, participants were told that 'you hired a temporary employee who usually works for another catering company—someone with experience, but under the understanding that at the end of the season he would be returning to his other job'.

### Dependent Variables

#### *Manipulation Check*

The effectiveness of the manipulation was checked by two items that assessed whether participants considered the offender to be an ingroup member: 'Do you consider the employee to be part of your group of employees?' (1 = *not at all*, 7 = *very much*) and 'Of all your employees, how central a group member is the employee in this situation?' (1 = *very peripheral*, 7 = *very central*). The two items were averaged ( $r = .64$ ).

#### *Justice Notions*

The main dependent measures were the same as in Study 2: A six item measure of restorative justice ( $\alpha = .80$ ) and a six-item measure of retributive justice ( $\alpha = .86$ ). The two scales were not significantly correlated ( $r = .08$ , ns). Exploratory factor analysis confirmed the distinction between the two justice notions, extracting two strong factors explaining 56% of the variance (a third factor with an eigenvalue greater than 1 was considerably weaker, and a Scree test provided evidence that a two-factor solution was more appropriate). All items loaded greater than .52 on their respective factor, with no substantial cross-loadings (see Appendix).

#### *Concerns*

As potential mediators, *value concerns* were measured by three items (e.g. The offender disregarded our common beliefs and values;  $\alpha = .79$ ), and *status/power concerns* also by three items (e.g. The incident was demeaning for me; 1 = *do not at all agree*, 7 = *strongly agree*;  $\alpha = .77$ ). The two concern scales were not significantly correlated ( $r = .10$ , ns). Exploratory factor analysis confirmed the distinction between the two concerns, extracting two factors with eigenvalues greater than 1 that explained 70% of the variance. All items loaded greater than .78 on their respective factor, with no substantial cross-loadings.

## Emotions

Using the stem 'How are you likely to feel?' following the offence, *sadness-related emotions* were measured by three items: Sadness, disappointment and shock (1 = *not at all*, 7 = *very much*;  $\alpha = .63$ ). *Anger-related emotions* were measured by another three items: Anger, contempt and disgust ( $\alpha = .68$ ; see Rozin, Lowery, Imada, & Haidt, 1999). The two emotion scales were moderately correlated ( $r = .22$ ,  $p = .008$ ). Exploratory factor analysis confirmed the distinction between the two emotions, extracting two factors with eigenvalues greater than 1 that explained 60% of the variance. All items loaded greater than .64 on their respective factor, with no substantial cross-loadings.<sup>1</sup>

## Controls

Measures of perceived *seriousness* of the offence ( $\alpha = .64$ ), the *intentionality* ( $\alpha = .82$ ) and *justifiability* of the offender's actions ( $\alpha = .73$ ) were the same as in Study 2.

## RESULTS

A *t*-test for the manipulation check confirmed that the group membership manipulation was effective,  $t(149) = 5.12$ ,  $p < .001$ ,  $d = .83$ . In the ingroup (permanent employee) condition, the offender was considered to be a more central member of the victim's ingroup than in the outgroup (temporary employee) condition ( $M_s = 5.00$  vs. 3.76). On the other hand, ingroup and outgroup conditions did not differ in the degree of seriousness attributed to the offence ( $M_s = 5.73$  vs. 5.66),  $t(148) = .45$ , ns; nor in its perceived intentionality ( $M_s = 5.38$  vs. 5.20),  $t(147) = .82$ , ns, or justifiability ( $M_s = 4.76$  vs. 4.55),  $t(150) = .95$ , ns. Generally, the participants perceived the offender's actions as clearly wrong and unacceptable.

In line with predictions, *t*-tests showed that the restorative justice notion was more strongly endorsed when the offender was an ingroup member than when he was an outgroup member ( $M_s = 5.52$  vs. 5.13),  $t(150) = 2.36$ ,  $p = .020$ ,  $d = .38$ . In contrast, retributive justice was less strongly endorsed when the offender was an ingroup than outgroup member ( $M_s = 3.67$  vs. 4.38),  $t(150) = -3.25$ ,  $p = .001$ ,  $d = .53$ .

Further *t*-tests were applied to investigate whether the identity manipulation affected the potential mediators: Concerns and emotions elicited by the offence. First, as anticipated, value concerns were marginally more pronounced when the offender was an ingroup than outgroup member ( $M_s = 5.80$  vs. 5.47),  $t(150) = 1.80$ ,  $p = .074$ ,  $d = .29$ . However, unexpectedly, the same was true for status/power concerns ( $M_s = 3.09$  vs. 2.68),  $t(150) = 1.76$ ,  $p = .080$ ,  $d = .29$ . Second, as expected, sadness-related emotions were significantly stronger when the offender was an ingroup than outgroup member ( $M_s = 5.47$  vs. 4.54),  $t(149) = 4.84$ ,  $p < .001$ ,  $d = .79$ . Anger-related emotions did not differ significantly between ingroup and outgroup conditions ( $M_s = 4.20$  vs. 4.32),  $t(149) = -.55$ , ns.<sup>2</sup>

## Mediation by Value and Status/Power Concerns

We first investigated whether the identity effects on justice notions were mediated by value and/or status/power concerns triggered by the transgression. First, for restorative justice, regression analyses showed, as already seen, (a) an ingroup (compared to outgroup) offender led to stronger endorsement of a restorative justice notion ( $\beta = .18$ ,  $p = .024$ ); and (b) an ingroup (compared to outgroup) offender elicited marginally stronger value concerns ( $\beta = .15$ ,  $p = .074$ ) and status/power concerns ( $\beta = .14$ ,  $p = .080$ ). Further, (c) regressing restorative justice simultaneously on both concerns while controlling

<sup>1</sup>We do not claim that sadness, disappointment and shock are interchangeable emotions; nor are anger, contempt and disgust the same (e.g. Rozin et al., 1999). However, in the present research these formed two distinct clusters of emotions, as anticipated.

<sup>2</sup>Considering Footnote 1, it may be of interest that the effects of group membership were similar for each of the three emotions included in the two emotion clusters. That is, disappointment, sadness and shock were all greater with an ingroup compared to outgroup offender ( $M_s = 6.28$  vs. 5.64; 4.41 vs. 3.42 and 5.72 vs. 4.54, respectively),  $t > 2.89$ . In contrast, anger, contempt and disgust did not differ between ingroup and outgroup conditions ( $M_s = 5.09$  vs. 5.11; 3.70 vs. 3.89 and 3.80 vs. 3.95, respectively),  $t < 1$ .

for the group membership manipulation, value concerns were significantly positively related to restorative justice ( $\beta = .44$ ,  $p < .001$ ), whereas status/power concerns were not ( $\beta = .04$ , ns). When controlling for value concerns, the effect of group membership on restorative justice was only marginally significant ( $\beta = .12$ ,  $p = .090$ ). A Sobel test showed that the indirect effect of group membership via value concerns was marginally significant ( $z = 1.73$ ,  $p = .088$ ), suggesting that the effect of shared identity on a restorative justice notion might be partially mediated by value concerns.

Second, for retributive justice, regression analyses showed, as already seen, an ingroup (compared to outgroup) offender led to weaker endorsement of a retributive justice notion ( $\beta = -.26$ ,  $p = .001$ ). However, when regressing retributive justice on both concerns while controlling for the group membership manipulation, both value concerns ( $\beta = .19$ ,  $p = .013$ ) and status/power concerns ( $\beta = .18$ ,  $p = .019$ ) were significantly *positively* related to retributive justice. Considering the positive effect of offender identity on value and status/power concerns, neither concern could thus mediate the negative effect of group membership. In fact, controlling for the two concerns, the effect of group membership became even more pronounced ( $\beta = -.31$ ,  $p < .001$ ).

### Mediation through Sadness- and Anger-Related Emotions

An equivalent analysis was run for emotions as potential mediators. First, for restorative justice, regression analyses showed, as already seen, (a) a total effect of group membership on the endorsement of a restorative justice notion ( $\beta = .18$ ,  $p = .024$ ); and (b) an ingroup (compared to outgroup) offender elicited significantly stronger sadness-related emotions ( $\beta = .37$ ,  $p < .001$ ), but no stronger anger-related emotions ( $\beta = -.04$ , ns). Further, (c) when regressing restorative justice simultaneously on both types of emotions while controlling for the group membership manipulation, sadness-related emotions were significantly positively related to restorative justice ( $\beta = .36$ ,  $p < .001$ ), whereas anger-related emotions were not ( $\beta = -.03$ , ns). When controlling for sadness-related emotions, the effect of group membership on restorative justice was no longer significant ( $\beta = .05$ , ns). A Sobel test showed that the indirect effect of group membership via sadness emotions was significant ( $z = 3.24$ ,  $p = .002$ ), indicating that the effect of shared identity on a restorative justice notion was fully mediated by sadness-related emotions.

Second, for retributive justice, regression analyses showed, as already seen, group membership had a negative effect on the endorsement of a retributive justice notion ( $\beta = -.26$ ,  $p = .001$ ). When regressing retributive justice on both types of emotions while controlling for the group membership manipulation, sadness-related emotions were not significantly related to retributive justice ( $\beta = -.11$ , ns), but anger-related emotions were ( $\beta = .52$ ,  $p < .001$ ). Considering the fact that the offender's group membership did not affect anger-related emotions significantly, the latter could not mediate the negative effect of shared group identity on retributive justice. In fact, controlling for anger emotions, the effect of group membership hardly changed ( $\beta = -.24$ ,  $p = .001$ ).

## DISCUSSION

Study 3 yielded three important insights. First, using yet another operationalization of shared identity, namely in terms of the nominal membership status of the offender vis-à-vis the victim's group, yielded once more a stronger endorsement of restorative justice when victim and offender had a shared identity. For ingroup offenders, victims sought a restoration of justice through reaffirming violated values by re-establishing consensus. Second, this finding contrasted with an opposite effect of offender group membership on the endorsement of retributive justice. The view that only punishment inflicted on the offender could restore justice was more strongly held for an outgroup offender. Different from Study 2, where the offender was always nominally included in the ingroup (only in differently cohesive ingroups), the offender's explicit outgroup status seems to have elicited a stronger endorsement of retributive justice.

Third, in line with our theoretical argument, it was found that the stronger endorsement of the restorative justice notion for ingroup offenders was fully mediated by greater feelings of sadness. There was also evidence, albeit less conclusive, that the transgression was more strongly appraised as violating and questioning shared values when the offender was an ingroup member; and this partially mediated a greater preference for a restorative justice notion. Again, these finding contrasted with results for the endorsement of retributive justice, which was related to feelings of anger (not sadness) and

concerns about status/power (but also values). Interestingly, however, neither anger-related emotions nor status/power concerns (or value concerns) mediated the effects of offender group membership on the endorsement of retributive justice. We will discuss this issue below.

## GENERAL DISCUSSION

Inspired by the relatively recent restorative justice movement in criminology and law (see Bazemore, 1998; Braithwaite, 1998), we distinguish conceptually between two psychological notions of justice restoration after a transgression: Retributive justice and restorative justice (Wenzel et al., 2008). Retributive justice is the restoration of justice through the imposition of punishment, in line with its common understanding as just desert: Due to the wrongdoing the offender deserves to be punished, and punishment restores the moral balance (Carlsmith et al., 2002; Feather, 1999). In contrast, we conceptualize restorative justice as the restoration of justice through the reaffirmation of values that have been violated but are expected to be shared between victim and offender (Okimoto et al., 2009; Wenzel & Thielmann, 2006). It implies a process of seeking consensus between the affected parties where the offender understands the extent of the harm done, accepts responsibility and offers an apology, while the victim acknowledges or affirms this new-found consensus, ideally by forgiving the offender.

The present research aims to further our understanding of when and why consensus-based, restorative justice might matter to victims and inform their responses to the victimization. We argue that victims might seek consensus in order to validate values that were violated by the offence. However, this reaffirmation of values should matter to them in particular when they expect to share the same values with the offender; only when they expect to share values with the offender would consensus contribute to the social validation of those values (Turner, 1987). The expectation of holding the same values, in turn, is based on the victims' perception that they share an identity with the offender that is relevant in the given situation (Turner, 1987). Hence, we argue that bilateral consensus should be particularly sought and valued when victims share a group membership with the offender and this social identity is salient and relevant to their self-definition (Wenzel et al., 2008). Across three different studies, two different paradigms and four different operationalizations of identity, we found evidence in line with our theoretical conceptualization of restorative justice.

Study 1 showed that an administrative justice response in a university context was considered fairer and more justice-restoring when it was based on a shared understanding between the parties. However, in line with our reasoning, this was the case in particular when the university identity, which participants shared with the offender, had been made salient or was chronically strong. Conversely, the mere, but *failed*, attempt to achieve an understanding led to greater perceptions of *unfairness* when participants identified strongly with their university. In Study 2, we measured more explicitly the endorsement of a restorative justice philosophy as a mediating concept. Participants endorsed restorative justice more strongly when the group they shared with the offender was cohesive and, thus, presumably more salient and identity-defining. In turn, the endorsement of restorative justice explained why victims, when sharing a relevant identity with the offender, showed a greater preference for actions that facilitate consensus, including demands that the offender accept accountability and offer an apology, as well as the victim's willingness to forgive. Similarly, in Study 3, using a different operationalization of shared identity, restorative justice was more strongly endorsed when the offender was an ingroup member than when he was an outgroup member.

Moreover, Study 3 provided evidence for potential mediating processes. When victims share an identity with offenders and thus expect them to hold the same values, they should be more likely to interpret the injustice as a violation and questioning of values. Hence, justice repair requires a reaffirmation of those values through renewed consensus. Likewise, when the transgression is interpreted primarily as a violation of shared values, feelings of sadness and disappointment should dominate; disappointment about the violation of an expectation, and sadness about the harm done to shared values and identity. The renewal of consensus should not only reaffirm the values but also alleviate the victim's emotional pain. Study 3 provided evidence for the mediating role of sadness-related emotions, which were stronger when the offender was an ingroup member and fully mediated the effects of shared identity on the adoption of restorative justice. The results are consistent with findings by Okimoto et al. (2009) where a consensus-based response was considered justice-restoring in particular when participants reported feelings of sadness in response to the injustice.

The findings for consensus-based, restorative justice contrast with the results for the traditional retributive justice notion. Retributive justice was not affected by the degree of group cohesiveness in Study 2, and it was in fact more strongly endorsed when the offender was an outgroup member in Study 3. We argue that this could be the case because victims might define themselves vis-à-vis an outgroup offender in more competitive terms. They thus might be more likely to interpret the transgression as a threat to their status or power, which should lead to frustration and anger, and the motivation to see justice restored through a reassertion of their status/power and a diminishing of the offender's status/power. However, while status/power concerns and anger emotions were related to retributive justice as predicted, they failed to mediate the effects of offender group membership.

It remains a question for future research what might mediate the effects of offender identity on retributive justice endorsement. One possibility is that concerns about certain aspects of one's status or power are in fact heightened in *intragroup* (shared identity) contexts, such as one's standing within a group or level of *respect*. This source of positive social identity needs to be distinguished from *intergroup* competition over status, power and resources, and the positive differentiation from relevant outgroups that is reflected in *pride* (see Tyler & Blader, 2001). Our current measures might not have differentiated between the two aspects or, indeed, might even have picked up more of an intragroup *respect* meaning of status/power (hence, in Study 3, stronger concerns were observed when the offender was an ingroup member). Future studies should include more differentiated measures that tap both aspects of status/power. We would predict that *pride* is the mediator driving the effects of offender outgroup identity on punitive, retributive justice. Consistent with this argument, van Prooijen and Lam (2007) recently observed that participants were more punitive towards an outgroup offender (but not an ingroup offender) when their own group's status was relatively low and, thus, their pride perhaps vulnerable. Note also that excessive personal pride, a narcissistic inflation of one's self-esteem, has been found to be a predictor of reactive aggression and revenge (Baumeister, Bushman, & Campbell, 2004).

It must be acknowledged that the current research focused on the distinctiveness of restorative and retributive justice. Measured as abstract one-dimensional philosophies, the two justice notions indeed proved uncorrelated. However, this also means that they are not mutually exclusive and, in practice, could be combined, raising important questions about their possible interplay (Gromet & Darley, 2006, 2009). For example, Okimoto and Wenzel (2009) have shown that punishment can also have a value-reaffirmation function, suggesting that the distinctive concern underlying restorative justice might not be about values *per se*, but rather how and with whom these are to be re-validated. Different justice responses, such as punishment, apology and forgiveness, may serve dual concerns (Okimoto & Wenzel, 2008), and they might do so in interaction with each other. This is another question for future research.

The major strength of the present research is the consistency of its finding across different methods and operationalizations of identity. However, there are some limitations. Specifically, Studies 2 and 3 used hypothetical scenarios with the instruction for participants to take the perspective of the victim. It is not clear how well the participants were able to put themselves into the victim's shoes, but obviously they were not as fully immersed in the situation as had they been real victims of such an incident. Research by Robinson and Clore (2001) indicates that scenario-based methodologies actually show remarkably good validity in terms of participants' appraisals and reported emotions as well as their interrelations. Further, Study 1 presented the transgression as a real incident with immediate relevance, as the data were collected under the pretence they would inform real university policy. However, while these methods might be good approximations given the ethical constraints of research in this field, future research should also attempt to investigate the validity of our findings with real victims. Alternative experimental approaches could stage the victimization of participants in the laboratory, or draw on real (or pseudo-real) events where participants are victims qua their membership in a group that has been subjected to an injustice. Correlational approaches could also draw on retrospective recall of experiences of victimization or, prospectively, record participants' experiences once they report a transgression against them.

While acknowledging the limitations, the present research adds evidence to the limited yet growing empirical support for the distinction between retributive and restorative justice notions. This conceptual distinction promises a better understanding of how and why people seek to restore justice following transgressions from a social psychological perspective. The victims' identity relation to the offender seems to be a crucial determinant of their justice notion. It may not simply be a question of whether people react more strongly to outgroup offenders because of their more favourable attitude to ingroup members, or whether they react more strongly to ingroup offenders because they regard the transgression as a betrayal of their ingroup (see Vidmar, 2002). Rather, in both cases, whether the offender is a member of the ingroup or outgroup, people will be motivated to restore justice—however, different forms of justice, for different reasons.

## ACKNOWLEDGEMENTS

This research was supported by a grant from the Australian Research Council, DP0557634.

## REFERENCES

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Aquino, K., Tripp, T. M., & Bies, R. J. (2006). Getting even or moving on? Power, procedural justice, and types of offense as predictors of revenge, forgiveness, reconciliation, and avoidance in organizations. *Journal of Applied Psychology, 91*, 653–668.
- Ashmore, R. D., Deaux, K., & McLaughlin-Volpe, T. (2004). An organizing framework for collective identity: Articulation and significance of multidimensionality. *Psychological Bulletin, 130*, 80–114.
- Averill, J. R. (1982). *Anger and aggression: An essay on emotion*. New York: Springer.
- Baumeister, R. F., Bushman, B. J., & Campbell, W. (2000). Self-esteem, Narcissism, and aggression: Does violence result from low self-esteem or from threatened egotism? *Current Directions in Psychological Science, 9*, 26–29.
- Bazemore, G. (1998). Restorative justice and earned redemption: Communities, victims, and offender reintegration. *American Behavioral Scientist, 41*, 768–813.
- Bies, R. J., & Tripp, T. M. (1996). Beyond distrust: ‘Getting even’ and the need for revenge. In R. M. Kramer, & T. R. Tyler (Eds.), *Trust in organizations: Frontiers of theory and research* (pp. 246–260). Thousand Oaks, CA: Sage.
- Braithwaite, J. (1998). Restorative justice. In M. Tonry (Ed.), *The handbook of crime and punishment* (pp. 323–344). Oxford: Oxford University Press.
- Brewer, M. B. (2007). The Social psychology of intergroup relations: Social categorization, ingroup bias, and outgroup prejudice. In A. W. Kruglanski, & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (2nd ed., pp. 696–715). New York: Guilford.
- Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science, 3*, 14–25.
- Carlsmith, K. M. (2006). The roles of retribution and utility in determining punishment. *Journal of Experimental Social Psychology, 42*, 437–451.
- Carlsmith, K. M., Darley, J. M., & Robinson, P. H. (2002). Why do we punish? Deterrence and just deserts as motives for punishment. *Journal of Personality and Social Psychology, 83*, 284–299.
- Castano, E., Yzerbyt, V., & Bourguignon, D. (2003). We are one and I like it: The impact of ingroup entitativity on ingroup identification. *European Journal of Social Psychology, 33*, 735–754.
- Christie, N. (1977). Conflicts as property. *British Journal of Criminology, 17*, 1–15.
- Cohen, R. L. (2001). Provocations of restorative justice. *Social Justice Research, 14*, 209–232.
- Daly, K. (2000). Revisiting the relationship between retributive and restorative justice. In H. Strang, & J. Braithwaite (Eds.), *Restorative justice: Philosophy to practice* (pp. 33–54). Aldershot: Ashgate/Dartmouth.
- Darley, J. M., Carlsmith, K. M., & Robinson, P. H. (2000). Incapacitation and just deserts as motives for punishment. *Law and Human Behavior, 24*, 659–683.
- Duff, R. A. (2001). *Punishment, communication, and community*. Oxford, UK: Oxford University Press.
- Durkheim, E. (1964). *The division of labor in society* (G. Simpson, Trans.). Glencoe, IL: Free Press. (original work published 1902).
- Exline, J. J., Worthington, E. L., Jr., Hill, P., & McCullough, M. E. (2003). Forgiveness and justice: A research agenda for social and personality psychology. *Personality and Social Psychology Review, 7*, 337–348.
- Feather, N. T. (1999). *Values, achievement, and justice: Studies in the psychology of deservingness*. New York: Kluwer/Plenum.
- Folger, R., Rosenfield, D., Grove, J., & Corkran, L. (1979). Effects of ‘voice’ and peer opinions on responses to inequity. *Journal of Personality and Social Psychology, 37*, 2253–2261.
- Govier, T. (2002). *Forgiveness and revenge*. London: Routledge.
- Graham, S., Weiner, B., & Zucker, G. S. (1997). An attributional analysis of punishment goals and public reactions to O. J. Simpson. *Personality and Social Psychology Bulletin, 23*, 331–346.
- Gromet, D. M., & Darley, J. M. (2006). Restoration and retribution: How including retributive components affects the acceptability of restorative justice processes. *Social Justice Research, 19*, 395–432.
- Gromet, D. M., & Darley, J. M. (2009). Retributive and restorative justice: The importance of crime severity and shared identity on people’s justice responses. *Australian Journal of Psychology, 61*, 50–57.
- Haslam, S. A., McGarty, C., & Turner, J. C. (1996). Salient group memberships and persuasion: The role of social identity in the validation of beliefs. In J. L. Nye, & A. M. Brower (Eds.), *What’s social about social cognition? Research on socially shared cognition in small groups* (pp. 29–56). Thousand Oaks, CA: Sage.
- Haslam, S. A., Oakes, P. J., Reynolds, K. J., & Turner, J. C. (1999). Social identity salience and the emergence of stereotype consensus. *Personality and Social Psychology Bulletin, 25*, 809–818.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Jetten, J., Branscombe, N. R., Spears, R., & McKimmie, B. M. (2003). Predicting the paths of peripherals: The interaction of identification and future possibilities. *Personality and Social Psychology Bulletin, 29*, 130–140.

- Jetten, J., Spears, R., & Manstead, A. S. (1997). Distinctiveness threat and prototypicality: Combined effects on intergroup discrimination and collective self-esteem. *European Journal of Social Psychology*, *27*, 635–657.
- Kahan, D. M. (1996). What do alternative sanctions mean? *University of Chicago Law Review*, *63*, 591–653.
- Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. T. Gilbert, & S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 233–265). New York: McGraw-Hill.
- Kerr, N. L., Hymes, R. W., Anderson, A. B., & Weathers, J. E. (1995). Defendant–juror similarity and mock juror judgments. *Law and Human Behavior*, *19*, 545–567.
- Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist*, *46*, 819–834.
- Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion*. New York: Plenum Press.
- Marques, J. M., Abrams, D., & Seródio, R. G. (2001). Being better by being right: Subjective group dynamics and derogation of in-group deviants when generic norms are undermined. *Journal of Personality and Social Psychology*, *81*, 436–447.
- Marques, J. M., & Paez, D. (1994). The ‘black sheep effect’: Social categorization, rejection of in-group deviates, and perception of group variability. In W. Stroebe, & M. Hewstone (Eds.), *European review of social psychology* (Vol. 5, pp. 37–68). Chichester, UK: Wiley.
- Marshall, T. F. (1998). *Restorative justice: An overview*. A report by the home office research development and statistics directorate. London: Home Office.
- Mikula, G. (2003). Testing an attribution-of-blame model. *European Journal of Social Psychology*, *33*, 793–811.
- Morrison, B., & Ahmed, E. (2006). Restorative justice and civil society: Emerging theory, practice and evidence. *Journal of Social Issues*, *62*, 209–215.
- Oakes, P. J. (1987). The salience of social categories. In J. C. Turner, M. A. Hogg, P. J. Oakes, S. D. Reicher, & M. S. Wetherell (Eds.), *Rediscovering the social group: A self-categorization theory* (pp. 117–141). Oxford, UK: Basil Blackwell.
- Ohbuchi, K.-I., Kameda, M., & Agarie, N. (1989). Apology as aggression control: Its role in mediating appraisal of and response to harm. *Journal of Personality and Social Psychology*, *56*, 219–227.
- Okimoto, T. G., & Wenzel, M. (2008). The symbolic meaning of transgressions: Towards a unifying framework of justice restoration. In K. Hegtvædt, & J. Clay-Warner (Eds.), *Advances in group processes* (pp. 291–326). Bingley, UK: Emerald.
- Okimoto, T. G., & Wenzel, M. (2009). Punishment as restoration of group and offender values following a transgression: Value consensus through symbolic labelling and offender reform. *European Journal of Social Psychology*, *39*, 346–367.
- Okimoto, T. G., Wenzel, M., & Feather, N. T. (2009). Beyond compensation and punishment: Conceptualizing restorative justice and exploring its determinants. *Social Justice Research*, *22*, 156–180.
- Robinson, M. D., & Clore, G. L. (2001). Simulation, scenarios, and emotional appraisal: Testing the convergence of real and imagined reactions to emotional stimuli. *Personality and Social Psychology Bulletin*, *27*, 1520–1532.
- Robinson, D. T., Smith-Lovin, L., & Tsoudis, O. (1994). Heinous crime or unfortunate accident? The effects of remorse on responses to mock criminal confessions. *Social Forces*, *73*, 175–190.
- Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, *76*, 574–586.
- Schubert, T. W., & Otten, S. (2002). Overlap of self, ingroup, and outgroup: Pictorial measures of self-categorization. *Self and Identity*, *1*, 353–376.
- Sommers, S. R., & Ellsworth, P. C. (2000). Race in the courtroom: Perceptions of guilt and dispositional attributions. *Personality and Social Psychology Bulletin*, *26*, 1367–1379.
- Tropp, L. R., & Wright, S. C. (2001). Ingroup identification as the inclusion of ingroup in the self. *Personality and Social Psychology Bulletin*, *27*, 585–600.
- Turner, J. C. (1981). The experimental social psychology of intergroup behaviour. In J. C. Turner, & H. Giles (Eds.), *Intergroup behaviour* (pp. 66–101). Oxford: Basil Blackwell.
- Turner, J. C. (1987). The analysis of social influence. In J. C. Turner, M. A. Hogg, P. J. Oakes, S. D. Reicher, & M. S. Wetherell (Eds.), *Rediscovering the social group: A self-categorization theory* (pp. 68–88). Oxford: Blackwell.
- Tyler, T. R., & Blader, S. L. (2001). Identity and cooperative behavior in groups. *Group Processes and Intergroup Relations*, *4*, 207–226.
- Van Prooijen, J.-W., & Lam, J. (2007). Retributive justice and social categorization: The perceived fairness of punishment depends on intergroup status. *European Journal of Social Psychology*, *37*, 1244–1255.
- Vidmar, N. (2000). Retribution and revenge. In J. Sanders, & V. L. Hamilton (Eds.), *Handbook of justice research in law* (pp. 31–63). New York: Kluwer/Plenum.
- Vidmar, N. (2002). Retributive justice: Its social context. In M. Ross, & D. T. Miller (Eds.), *The justice motive in everyday life* (pp. 291–313). New York: Cambridge University Press.
- Walster, E., Walster, G. W., & Berscheid, E. (1978). *Equity: Theory and research*. Boston: Allyn And Bacon.
- Wenzel, M. (2002). What is social about justice? Inclusive identity and group values as the basis of the justice motive. *Journal of Experimental Social Psychology*, *38*, 205–218.
- Wenzel, M., Okimoto, T. G., Feather, N. T., & Platow, M. J. (2008). Retributive and restorative justice. *Law and Human Behavior*, *32*, 375–389.
- Wenzel, M., & Thielmann, I. (2006). Why we punish in the name of justice: Just desert versus value restoration and the role of social identity. *Social Justice Research*, *19*, 450–470.
- Worthington, E. L. jr., (Ed.). (2005). *Handbook of forgiveness*. New York: Routledge.

## APPENDIX

Levels of endorsement of restorative and retributive justice notions were measured by six items each (1 = *do not at all agree*, 7 = *strongly agree*). The items and results for factor analyses in Studies 1 and 2 are presented in Table A1.

Table A1. Factor loadings for restorative and retributive justice items (varimax rotation)

	Study 2		Study 3	
	Factor 1	Factor 2	Factor 1	Factor 2
As a matter of fairness, the employee should be penalized	<b>.77</b>	.29	<b>.60</b>	.00
The only way to restore justice is to punish the employee	<b>.87</b>	.03	<b>.84</b>	.09
Justice is served at the moment that the employee is punished	<b>.71</b>	-.08	<b>.80</b>	.00
Only a punishment restores the justice disrupted by the incident	<b>.78</b>	.00	<b>.84</b>	.01
For the sake of justice, some degree of suffering has to be inflicted on the employee	<b>.79</b>	-.10	<b>.74</b>	.06
The employee deserves to be penalized	<b>.80</b>	.01	<b>.80</b>	.04
For justice to be reinstated, the affected parties need to achieve agreement about the values violated by the incident	.09	<b>.73</b>	.11	<b>.73</b>
To restore justice, the employee and I need to reaffirm consensus on the values and rules in a workplace	-.27	<b>.70</b>	-.04	<b>.78</b>
Without the employee's sincere acknowledgement of having acted inappropriately, the injustice is not completely restored	.21	<b>.66</b>	.05	<b>.53</b>
A sense of justice requires that the employee and I develop a shared understanding of the harm done by the incident	-.11	<b>.76</b>	-.04	<b>.78</b>
Justice is restored as soon as the employee has learnt to endorse the values violated by the incident	.10	<b>.58</b>	.00	<b>.69</b>
For a sense of justice, we all, including the employee and I, need to reaffirm our belief in shared workplace values	-.02	<b>.71</b>	.07	<b>.76</b>