University of Groningen

Was it One of Us?

Otten, Sabine; Gordijn, Ernestine

Published in:
Social and Personality Psychology Compass

DOI:
[10.1111/spc3.12098](https://doi.org/10.1111/spc3.12098)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2014

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Otten, S., & Gordijn, E. (2014). Was it One of Us? How People Cope with Misconduct by Fellow Ingroup Members. *Social and Personality Psychology Compass*, 8(4), 165–177. <https://doi.org/10.1111/spc3.12098>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Was It One of Us? How People Cope with Misconduct by Fellow Ingroup Members

Sabine Otten* and Ernestine H. Gordijn

University of Groningen

Abstract

How do group members cope with misconduct by members of their own group? Strong evidence for a positive bias in people's evaluations of their own group and its members suggests that an ingroup perpetrator is likely to be treated more leniently than an outgroup perpetrator. However, research has also demonstrated a "Black Sheep-Effect", such that ingroup members who deviate negatively from relevant ingroup norms are evaluated and punished even more harshly than outgroup deviants. Both reactions – positive ingroup bias and the Black Sheep Effect – may serve the same goal, namely maintaining positive regard for the ingroup. In this paper, we present several moderators that have been shown to affect responses to negative ingroup deviance. We propose a model that incorporates and organizes these moderators in order to predict whether negatively deviant ingroup members will receive especially lenient or rather harsh and negative evaluations from their fellow ingroup members.

When a 16-year-old Dutch girl named Marianne Vastra was found raped and murdered in a meadow in 1999, people were quick to suggest that the murderer was likely to be found in the refugee center near the village where the girl was killed. Indeed, many people in the Netherlands were quite convinced that only someone from abroad could commit such a horrific murder. However, in February 2013, when the murderer was identified as a local farmer, people were shocked that the crime was committed by a member of their community. Some people, including the mother of the victim, refused to believe that it was true and criticized the DNA methods that were used to solve the crime. Others, however, complained that the sentences in the Netherlands were not harsh enough to punish the (ingroup) murderer.

'Was it one of us?' is a question that often arises when people are confronted with such atrocious behavior. As people are inclined to think positively about themselves and the groups they belong to (e.g., Baumeister, Tice, & Hutton, 1989; Hewstone, Rubin, & Willis, 2002), they are likely to assume that only members of other groups could possibly be responsible. The question is: what happens when people learn that this hypothesis was wrong and it was actually 'one of us'. On the one hand, a plausible reaction could be finding excuses for the ingroup perpetrator or downplaying the severity of the event. However, as our opening example shows, the opposite reaction is also likely: Ingroup perpetrators are sometimes rejected and punished even more harshly than perpetrators from other groups (i.e., outgroups) engaging in the same misconduct.

As we will demonstrate in this article, both types of reactions – downplaying or questioning the ingroup perpetrator's misconduct versus especially harsh evaluations and punishment – indeed occur. Moreover, we assume that both reactions are driven by the same underlying motive: a striving for a positive social identity. Yet, while this motive is stable, the way in which it is strived for will vary as a function of several moderating variables. In the present article, we will introduce a model of Coping with Ingroup Deviance (CID) that

organizes existing evidence on how people deal with negative ingroup deviance, a model that may also inspire new research in this field. Importantly, we argue that when relevant moderators are taken into account, it can be predicted whether people will show very mild or rather very harsh reactions when dealing with ingroup deviance.

Social Categorization, Ingroup Favoritism, and Coping with Negative Ingroup Deviance

People have a tendency to categorize their social world into broad categories, for example, into men versus women or children versus adults or Christians versus Muslims. A very basic yet important distinction that we make when categorizing our social world is between ingroups (“us”), to which we belong, and outgroups (“them”), to which we do not belong. According to Social Identity Theory (Tajfel & Turner, 1979), we derive part of our identity from our membership in social groups. For example, who we are is, amongst other things, determined by our nationality, our gender, and/or our professional affiliation. Just as people tend to strive for positive self-regard (e.g., Baumeister, 1998; Baumeister et al., 1989), they will also strive for positive evaluations of their group memberships. Such positive social identities are often achieved by perceiving the ingroup as comparatively better than other groups. This ingroup favoritism has been found in many empirical studies and for a broad variety of groups (e.g., Brewer, 1979; Hewstone, Rubin, & Willis, 2002; Mullen, Brown, & Smith, 1992).

Although the positive bias towards ingroups is robust, it is *not* unconditional; instead, it has to be responsive to *reality constraints* (for overviews, see Brewer & Brown, 1998; Hewstone et al., 2002). For example, soccer supporters whose team just lost miserably in competition with another team will, in the short run, not be prone to brag about the superior quality of their team’s soccer performance. They will, however, likely try to reduce the damage of their own group’s defeat, for example, by interpreting their team’s bad performance as only temporary (e.g., due to core player injuries). As we will further elaborate in this article, this responsiveness of ingroup bias to reality constraints is crucial for understanding differences in *how* people cope with negative ingroup deviance.

Positive bias towards ingroup deviants

Considering group member’s interest in a positive social identity, it makes sense to assume that people tend to be positively biased when confronting an ingroup member who shows unacceptable behavior. Group members will be motivated to take the perspective of such an ingroup deviant as well as give the benefit of the doubt if possible. As a consequence, they may downplay the harm that is done or make situational attributions (e.g., ‘he had trouble at work and therefore lost his temper’) rather than assuming that harm was done intentionally. In this way, the misbehaving ingroup member does not endanger the positive perception one has of one’s ingroup, as one is able to find excuses for the deviant behavior.

Moreover, a biased interpretation of the negative event need not require an extensive effort on the part of the perceiver. Research has shown that people have an automatic tendency to perceive their own groups positively. Such a general positive default towards the ingroup can spontaneously and unconsciously affect evaluations of the ingroup and its members (see, e.g., Otten & Moskowitz, 2000; Otten & Wentura, 1999). If ingroup members are seen positively at the outset, then such a default may override (or at least tone down) the potential negative evaluations that might otherwise occur when ingroup members violate norms.

Focusing on ‘real’ instances of intergroup violence between Catholics and Protestants in Northern Ireland, Hunter, Stringer, and Watson (1991) have provided evidence for such

ingroup biases in the interpretation of violent norm violations. Violence enacted by ingroup members was predominantly attributed to situational, external factors; violence enacted by outgroup members, however, was predominantly attributed to dispositional (personality-related) factors. Through this pattern of attributions, ingroup violence can be conceptualized in terms of unusual circumstances, while outgroup violence is a matter of hostile people with bad characters (for a review of more relevant evidence, see Hewstone, 1990).

Negative bias towards ingroup deviants

Besides a positively biased interpretation of ingroup misconduct and its causes, another reaction to deviant members within the ingroup is also possible and might be equally functional in maintaining positive views of the ingroup: Group members could react very negatively to the misbehaving ingroup member and may respond even *more* negatively than towards outgroup members performing the same behavior. By doing so, they can demonstrate that the deviant's behavior departs strongly from the norms of the group. This type of reaction to ingroup deviance was first demonstrated by Marques, Yzerbyt, and Leyens in 1988 and labeled the "*Black Sheep Effect*" (also see Marques & Yzerbyt, 1988).

Marques, Abrams, Paez, and Martinez-Taboada (1998) showed that people, who were portrayed as dislikable and norm-deviant, received more disapproval when categorized as an ingroup rather than outgroup member. Importantly, Marques, Yzerbyt, and Leyens (1988) argued that the Black Sheep Effect can be understood as a sort of 'sophisticated ingroup favoritism' (Marques et al., 1988, p. 5): By strongly distancing the average ingroup member from the ingroup deviant, and harshly criticizing the deviant's conduct, group members can act as if the deviant is not typical for the group and that the deviant behavior is an inappropriate exception to the proper conduct of other group members. This idea was further elaborated in the Subjective Group Dynamics Model (Marques, Abrams, & Serôdio, 2001). According to this model, negative reactions to ingroup deviants are functional because one can retain the perception that the group as a whole is benevolent despite a bad apple in the bunch (also see Marques et al., 1998). Imagine, for example, that the media finds out that a local politician from the Green Party (with a program placing a strong emphasis on protecting the environment) drives a gas-guzzling Jaguar to his appointments. By seeing the politician's behavior as in no way representative of the group, the party can reaffirm its commitment to environmentally friendly behavior and continue to feel positively about the party relative to other political parties. In short, an integrative model is needed to capture these cases versus those in which the deviant benefits from being an ingroup member.

An integrative model on coping with negative ingroup deviance

Together, these arguments imply that a striving for positive ingroup distinctiveness, as postulated in Social Identity Theory (Tajfel & Turner, 1979), makes – in principle – both options equally plausible and functional – i.e., a mild excusing versus a harsh rejection of ingroup deviants. The question becomes the circumstances under which ingroup deviants will be viewed more positively or more negatively than outgroup deviants. We believe that there are three important classes of moderating variables here, respectively involving (i) characteristics of the misconduct, (ii) characteristics of the group member committing the misconduct, and (iii) characteristics of the group member evaluating the misconduct. These three elements can be considered key aspects of the situation: There is a *perpetrator* or actor who has engaged in a *behavior* that negatively deviates from relevant norms, whose conduct will

elicit a response from a *victim* or an *observer*. Aggressive behaviors, regardless of this ingroup/outgroup context, are evaluated with respect to the same sorts of factors (e.g., Mummendey & Otten, 1993). Therefore, the factors should also be useful ones in thinking about how negatively deviant ingroup behaviors are evaluated.

We have summarized this reasoning in our Coping with Ingroup Deviance (CID) model (see Figure 1). In the remainder of this article, we will use this model to organize and discuss existing evidence on positively versus negatively biased reactions to ingroup perpetrators. Although the review captures the present state of affairs in this field of research, it will not be fully exhaustive. Within each group of moderators, for example, additional variables might be identified that would further refine our understanding of how people cope with ingroup deviance.

In this review, we will mostly consider evaluations and punishment recommendations. Where applicable, we will also refer to evidence on the specific emotions elicited by ingroup deviant behavior because different emotions are likely to prompt different behavioral reactions (Carver & Harmon-Jones, 2009; Frijda, 1986). Moreover, although we will concentrate on how people cope with misconduct by ingroup members, comparisons related to outgroup deviance will be made when informative evidence is available.

Figure 1. The Coping with Ingroup Deviance (CID) model.

Coping with Negative Ingroup Deviance: Relevant Moderators

In the following, we will discuss the three groups of moderators in some detail. *Characteristics of the misconduct* include considerations of which norms are violated and whether intent is ambiguous. *Characteristics of the group member* include the person's relative longevity and prominence in the group. *Characteristics of the evaluator* include variations in ingroup identification and cognitive capacity.

Characteristics of the misconduct

Violating generic versus ingroup-specific norms. When introducing the Black Sheep Effect, Marques et al. (1988) manipulated whether the deviant group member was offending either a general norm (applying to everyone) or a norm that particularly characterized the ingroup. The findings from this initial study revealed that only the ingroup-defining norm deviation elicited the Black Sheep Effect. However, in later research, it was found that offenses to generic norms can also sometimes elicit strong derogation of ingroup members (Marques et al., 2001). Given the latter research, it is probably safe to say that both distinguishing group norms (e.g., not eating pork for Muslims) and more generic humanity norms (e.g., not stealing) can produce the conditions necessary for the Black Sheep Effect.

Ambiguity of intent. Once a perpetrator is clearly identified, a determinant of reactions is whether the respective behavior (and its negative outcome) was intended or not (e.g., Mummendey, Linneweber, & Löschper, 1984). Indeed, it may be that as long as it is not unequivocally determined that a norm violation was actually intended, one can continue to feel good about the ingroup without harsh evaluations of the person engaging in the misconduct. Hence, the Black Sheep Effect may be less likely to occur if intent is ambiguous.

Research by Braun, Otten, and Gordijn (2009a; also Braun, 2010) supports this idea. The researchers manipulated the ambiguity of the intention to do harm and the categorization of the actor as either an ingroup or outgroup member in two studies. In study 1, a computer-simulated interaction paradigm was used. To win a prize, female participants needed the help of a virtual interaction partner in a simple problem-solving task. The interaction partner was introduced as either male (outgroup) or female (ingroup). In the ambiguous condition, a short video fragment showed the alleged interaction partner behaving in a way that might be an indication of unhelpfulness. In the unambiguous condition, the person could be heard saying that (s)he was unwilling to help. The results revealed that if the intention to do harm was ambiguous, an ingroup (i.e., female) non-helper was rated less negatively than an outgroup non-helper. But when intent was unambiguous, the pattern reversed. Now an ingroup member refusing to help was viewed particularly negatively.

Another way of characterizing the results of Braun et al. (2009a) is to say that the ambiguity manipulation had a much stronger effect if the perpetrator was of the ingroup rather than outgroup. A similar pattern was evident in the amount of *anger* participants reported (see Figure 2) and in the amount of *retaliatory aggression participants engaged in* (and measured via an adaptation of the hot-sauce-paradigm: Lieberman, Solomon, Greenberg, & McGregor, 1999). Moreover, anger played a mediating role. Specifically, the aggression exhibited against ingroup members intending their unhelpfulness was due to greater anger in this cell of the design relative to other cells.

In a second study (Braun et al., 2009a), student status was used as an intergroup context. Participants, who were all students, had to imagine a scenario in a crowded bar in which a

Figure 2. Effect of ambiguity of hostile intention and social categorization on anger about the misconduct; 1 = not at all angry, 5 = very angry (adapted from Braun, 2010, study 2.1).

student or non-student bumped into them, in turn causing them to spill their wine all over themselves. In the ambiguous condition, it was not clear whether the other patron had accidentally stumbled or the bump was deliberate. In the unambiguous condition, by contrast, the intent to push the participant was clearly present. Again, results revealed that attributions of intent figured strongly in how ingroup members were evaluated: Reactions were mild if intent was ambiguous and especially harsh if intent was clear. By contrast, such perceptions of intent were not important to how outgroup members were judged.

Van Prooijen (2006) examined similar processes. In this case, ambiguity related to whether one could be sure that a target had committed the misbehavior. In two experiments, van Prooijen found that when guilt was certain, feelings of anger and hostility were stronger towards an ingroup suspect than an outgroup suspect. However, if guilt was uncertain, the pattern reversed, such that the ingroup suspect aroused less anger and hostility than the outgroup suspect. Moreover, in a third experiment, the same pattern was shown with punishment intentions: As long as guilt was uncertain, the ingroup suspect (to a greater extent than the outgroup suspect) was given the benefit of the doubt with respect to recommended punishment.

In sum, this research reveals that when evaluating ingroup members, perceivers are attentive to cues to intent ambiguity. If intentions to do harm are unambiguous, then reality does not allow for a positive bias towards ingroup deviants. Rather, especially negative evaluations appear to serve the purpose of distancing most of one's group from the perpetrator. Such processes seem less important if outgroup members are involved. Finally, there is evidence that anger is a relevant mediator of evaluative and behavioral response to ingroup deviants (Braun, 2010; Braun et al., 2009a).

Characteristics of the person committing the misconduct

Newcomers versus established group members. Another relevant moderator concerns the membership status of the deviant. Pinto and collaborators (Pinto, Marques, Levine, & Abrams 2010) distinguished three types of group members: First, there are *newcomers* who have just entered

the group and are in the socialization phase of their membership. Second, there are *full members* who have successfully completed socialization, which gives them both more privileges and more responsibilities for the group and its goal achievements. Third, there are *marginal members* who lost their full member status by not fulfilling group standards. These members may either re-socialize, in which case they can be considered ingroup members again, or they may psychologically exit (or be exiled by) the group.

Pinto et al. (2010) argued that full members are particularly representative of the group and its positive image. Therefore, these full members should be evaluated most strictly when they do wrong, and they should receive the most praise when they engage in normative behavior. To examine these ideas, the researchers presented participants with descriptions of fellow ingroup members (students) who either expressed opinions in line with relevant group norms (e.g., “university students should join together and ... struggle for a better educational system”) or opinions conflicting with these norms (“university students are not mature enough to know what is best for them”). In addition, the target person was described as either a new group member (having studied for 6 months) or a full member (having studied for 4 years).

In support of their ideas, Pinto et al. (2010) found that evaluations (e.g., as selfish versus altruistic) of the deviant ingroup member were more negative when the target was a full rather than new member. Moreover, in two of the three experiments, a third condition was realized, one in which the target person was described as somewhat marginal (e.g., a student who had studied for several years but might not stay at school). Such marginal targets were evaluated as moderately as deviant new group members and thus more favorably than full member deviants. The authors also examined behaviors that would be directed toward the target. If the target was a new member, the goal would be to get the target to change his/her opinions. If the target was a full member, on the other hand, people desired to mete out punishment. Neither of these two reactions was especially prominent if the group member was a marginal one.

In sum, Pinto et al. (2010, p. 116) concluded that group members have a strong interest in the maintenance of norms that contribute to the ingroup's positive identity. As a consequence of such motives, they will perceive negatively deviant behavior by fellow ingroup members as threatening and react accordingly. This will be less true, however, if the deviant does not possess (perceived) full membership status.

Leaders versus ‘regular’ group members. Besides distinguishing between new and established group members, another relevant aspect that may affect reactions to ingroup deviance is the power that the deviant holds within the group. Abrams and collaborators (Abrams, Randsley de Moura, Marques, & Hutchison, 2008) argued that two opposing hypotheses can be made with respect to evaluations of a powerful (versus regular) deviant group member. On the one hand, the leader typically earns his or her status for being prototypical (Haslam, 2001; Hogg, 2001). Consequently, anti-normative behavior should strongly violate expectations and be evaluated more negatively (Biernat, Vescio, & Billings, 1999). On the other hand, as power is associated with leadership status, the leader “has permission to deviate” and “the leadership role may confer the potential to define the prescriptive norm” (Abrams et al., 2008, p. 664).

Abrams et al. (2008) argued that which of these two phenomena will apply may depend on the *phase of leadership* the deviant group member is in. They distinguish between future, current, and past leaders. For future leaders, the conferral criterion should apply: The group should give them “innovation credit”, providing them with permission to deviate from what is currently normative. Norm-deviant current and past leaders, however, will more probably be evaluated in terms of their failure to represent group expectations, which should result in

comparatively more negative appraisals. Additionally, as ex-leaders are no longer in the center of attention, they should elicit weaker negative reactions than current leaders. These expectations were broadly supported in a series of experiments. Past leaders expressing attitudes that were opposing the group norm were evaluated less negatively than current leaders, and the greatest deviance tolerance occurred for future leaders.

In two studies in our own lab (Otten & Gordijn, 2008), we also focused on reactions to negative deviance by powerful ingroup [as powerful ingroup (versus outgroup)] members while comparing currently powerful deviants with those who were powerful in the past, but no longer. The focus of deviance was different than in Abrams et al. (2008). In our research, norm violations offended general laws such as those prohibiting sexual harassment. The negative behaviors were ascribed to either an ingroup or outgroup member, but guilt was not totally unambiguous. In study 1, we referred to fictitious leaders who were either Austrian (outgroup) or Dutch (ingroup), and who were accused of having used drugs at work. In study 2, we took advantage of a realistic example case: the sexual harassment charge against former Dutch Prime Minister Lubbers. Lubbers was portrayed either as being Dutch (ingroup) or as being Christian Democrat (a conservative political party that is typically not voted for by Dutch psychology students and hence an outgroup). Further, the emphasis was put either on the fact that the leader was not powerful anymore or on the fact that he still had a powerful position.

In both studies, and in line with Abrams et al. (2008), we (Otten & Gordijn, 2008) found that past powerful ingroup members were evaluated more mildly than currently powerful ingroup members. Similarly, more guilt, shame, and anger as well as greater punishment intentions were found in response to powerful deviant ingroup members than in response to past leaders from the ingroup. Interestingly, the reverse pattern emerged for outgroup deviants (see Figure 3).

In sum, data from our (Otten & Gordijn, 2008) two studies showed that a currently powerful ingroup leader elicits decidedly negative evaluative and emotional reactions, even when there is some ambiguity with respect to guilt. These reactions were as (if not more) negative than those elicited by a deviant outgroup leader, thereby resembling the Black Sheep Effect. However, when emphasizing a loss of power, reactions change dramatically.

Figure 3. Effect of power and social categorization on perceived inappropriateness of misconduct; 1 = not at all inappropriate, 9 = very inappropriate (Otten & Gordijn, 2008; study 1).

Now we find lenience. Altogether, for ingroup members in powerful positions, the “benefit of the doubt” normally given to ingroup members is sometimes omitted due to their importance to the group as a whole.

A recent publication by Karelaia and Keck (2013) provided further nuances regarding the role of a perpetrator's leadership status in reactions to his or her misconduct. These authors found that when the leader's anti-normative behavior was severe (e.g., sexual harassment, verbally abusing a co-worker), (s)he elicited harsher reactions (i.e., stronger punishment intentions) than non-leaders performing the same behavior. However, the reverse pattern emerged when the severity of anti-normative behavior was rather minor (e.g., being late for a meeting, making personal phone calls at work). With such minor misbehaviors, there was greater leniency for the leader than non-leader. Moreover, while the impact of leadership status in the case of severe negative deviance was mediated by perceived betrayal, in the case of minor negative deviance, perceived entitlement was the relevant mediating variable.

Finally, recent research by Abrams, Randsley de Moura, and Travaglino (2013) further corroborates that – under certain conditions – leaders within the group might be treated more leniently than regular group members showing the same negatively deviant behavior. These authors argue that as long as a leader's behavior is driven by the goal to serve the group's interest, (s)he profits from a “transgression credit”. To put it differently, the assumption is that a “double standard” is applied when evaluating norm deviations committed by leaders compared with regular group members. What is considered as justified by the necessities of the situation in the former case (for a leader) might be disapproved of in the latter case. This assumption was tested among participants involved in sports. Reactions to scenarios describing transgressions (i.e., arguing vehemently with the referee and acting offensively towards opposing players) differed as a function of the deviant's role within a group: Evaluations were significantly more positive if a captain rather than a regular team member was the transgressor. These influences of leader status can also interact with characteristics of the misconduct (e.g., its severity). See Abrams et al. (2013) and Karelaia and Keck (2013) for further details.

Interestingly, both studies, the one by Abrams et al. (2013) and by Karelaia and Keck (2013) not only show that characteristics of the person showing the misconduct matter but also that these can interact with characteristics of the misconduct (e.g., the severity).

Characteristics of the person evaluating the misconduct

Ingroup identification. In early publications on the Black Sheep Effect, the term ‘ingroup identification’ was used to refer to mere membership (or non-membership) in the deviant's group (e.g., Marques, Robalo, & Rocha, 1992; Marques et al., 1988). Sometimes, ingroup identification has been further treated as a manipulation check (e.g., Pinto et al., 2010): At least moderate levels of identification were considered a necessary condition to properly distinguish between ingroup and outgroup deviance. Even though this reasoning is plausible, it does not sufficiently do justice to the potential relevance of group identification.

The point here is that the *strength* with which one identifies with an ingroup should matter (Marques et al., 1998). For example, unambiguous anti-normative behaviors exhibited by ingroup members should be more threatening as identification with the group increases. Evaluations of deviants should be harsher as a result of this greater level of group identification. Braun et al. (2009a) found just such a result: When hostile intent was unambiguous, those who identified with the ingroup to a greater extent gave the most negative evaluations

and made strongest punishment recommendations. Accordingly, we advocate further research in which variations in ingroup identification are treated as predictors of reactions and not just as manipulation checks.

Cognitive capacity. As mentioned earlier, there is evidence for the idea that the ingroup bias can be viewed as sort of default reaction, operating automatically to protect ingroup positivity (e.g., Amodio & Mendoza, 2010; Otten & Moskowitz, 2000; Otten & Wentura, 1999). Hence, categorizing a person as an ingroup member should almost immediately elicit a motivated tendency to view the person with benevolence. If so, how can we account for the ‘sophisticated ingroup favoritism’ (Marques et al., 1988, p. 8) of the Black Sheep Effect? Is this more negative evaluation of deviant ingroup members automatic, or does it rely on more effortful information processing?

Such questions concerning the automaticity of the Black Sheep Effect were investigated by our research group (Braun, Otten, & Gordijn, 2009b). In two scenario-based studies, we manipulated whether an aggressively acting ingroup member could be unambiguously identified as acting upon hostile intentions or not. As reviewed earlier, people tend to give ingroup members the “benefit of the doubt” when intent is ambiguous but decidedly not when hostile intent is clear. Of more importance for present purposes, we manipulated information processing capacity by introducing a high cognitive load in one condition but not the other. Cognitive load was established by having participants memorize a nine-digit number (Gilbert & Hixon, 1991) while reading and evaluating the scenario describing an ingroup member’s negative deviance.

Both of the studies (Braun et al., 2009b) suggested that a harsh derogation of unambiguously negative ingroup deviance is a strategic response that requires cognitive resources: In the condition without cognitive load, this deviant was seen as having acted inappropriately, and people felt angry in this condition. However, when cognitive resources were depleted, appraisals of inappropriateness and feelings of anger were not only reduced but they were also not influenced by the intent ambiguity manipulation (see Figure 4). Hence, these findings suggest that derogating deviant ingroup members requires cognitive effort. Or, stated in

Figure 4. Effect of cognitive capacity and ambiguity of hostile intention on the perceived inappropriateness of ingroup deviance; 1 = not at all inappropriate, 7 = very inappropriate (adapted from Braun, 2010, study 3.1).

other terms, it is likely a strategic, deliberate response in situations when the ingroup's reputation is endangered and positive expectations about ingroup members are challenged.

Conclusions

How do people react to behavior by fellow ingroup members that offends norms and thereby endangers the positive reputation of the ingroup? Our review of empirical findings shows that this coping is to a large extent guided by a positive bias toward the ingroup and its members. This bias, however, can take two forms – a traditional one in which the misbehavior is excused or downsized, or a more sophisticated one in which the deviant ingroup member receives especially harsh reactions that designate his or her status as black sheep (within a flock of otherwise white ones).

We have offered a model of Coping with Ingroup Deviance (CID) that can be used as a tool to better predict which form reactions to ingroup deviance will take. This model specifies several moderators of reactions to ingroup deviance, moderators grouped into the categories of characteristics of the misconduct, characteristics of the person committing the misconduct, and characteristics of the person evaluating the misconduct. Although we have mostly focused on main effects, the moderators should also interact with each other. Indeed, research by Karelaia and Keck (2013) and Abrams et al. (2013) begins to point to some of these interactive effects, thus far primarily with respect to the first two sets of CID moderators.

Our overview of findings from our own and other labs suggests that sharply distancing from the ingroup deviant occurs if reality constraints do not allow for mild disapproval (for example, because the norm deviation is strong or hostile intentions are unambiguous). In addition, a deviant's status appears to play an important role. For example, different normative constraints seem to apply to newcomers and leaders when compared to regular group members. Even when other moderator variables favor the Black Sheep Effect, though, it appears that the effect is contingent on some degree of processing capacity (Braun et al., 2009b). This may imply that positive ingroup feelings represent a default that can be corrected only when group members engage in systematic processing of the available evidence.

In some of the studies reported in this review, emotions as well as evaluations were assessed (e.g., Braun et al., 2009a, 2009b; Otten & Gordijn, 2008; Van Prooijen, 2006). Our CID model appears relevant in understanding these emotional reactions to deviance. Yet, rather than viewing emotions as simply another set of dependent measures, emotions may be more profitably viewed as mediators of relations between appraisals of the deviant act and behavioral responses such as punishment. For example, Biernat and collaborators (1999) have argued that the Black Sheep Effect may be due to expectancy disconfirmation. As expectations regarding ingroup members are typically more positive than those regarding outgroup members, stronger feelings of disappointment, and therefore stronger behavioral reactions, may occur when deviants are ingroup members. A proper investigation of the role of emotions in reaction to deviance, however, may necessitate moving away from scenario-based research to settings in which stronger emotional experiences are more likely to occur.

Obviously, our list of possible moderators of reactions to deviance is not complete. For example, the context in which the deviance occurs (e.g., in front of an ingroup or an outgroup audience) could be relevant, as could the question of whether the misconduct is chronic or represents a first offense. People may be more lenient towards ingroup members in the latter case. Yet, all findings reported in this article are highly consistent with the assumption that people's coping with ingroup deviance is guided by the motivation to maintain or reestablish the ingroup's positive regard. While this goal can be considered stable, the specific nature and circumstances of the misconduct will determine how it is fulfilled.

Short Biographies

Sabine Otten is a Full Professor at the Social Psychology Department at the University of Groningen, The Netherlands. She obtained her PhD in Psychology in 1991, at the University of Muenster, Germany. Before moving to the University of Groningen in 2002, she held positions at the University of Muenster and at the University of Jena, Germany. Dr Otten's research interests relate to the fields of intra- and intergroup relations and social integration. In particular, she is or has been working on the (cognitive) determinants of ingroup favoritism and ingroup identification; on the determinants of interpersonal, intra- and intergroup aggression; and, most recently, on cultural diversity in organizations. Dr Otten has been associate editor of the *European Journal of Social Psychology* and is or has been consulting editor for *Journal of Personality and Social Psychology*, *Personality and Social Psychology Bulletin*, and *Personality and Social Psychology Review*. In 2008, she was elected as member of the Executive Committee of the European Association of Social Psychology.

Ernestine Gordijn is Full Professor at the Social Psychology Department at the University of Groningen, The Netherlands. She obtained her PhD in 1998 from the University of Amsterdam, The Netherlands. Since 2001, she is working at the University of Groningen, the Netherlands. Her research interests refer to two main questions: How do people form impressions of others, and how do people expect to be perceived by other people? What consequences does this have for how people feel, think, and behave in social interaction? For example, Dr Gordijn studies how the impact of meta-stereotypes (i.e., expectations about how the own group will be seen by members of other groups) will influence emotions, cognitions, and behavior. Other relevant themes are social rejection, intergroup conflict, power relations, humility and humiliation in interpersonal and intergroup relations, charisma, and minority influence. Dr Gordijn is currently editor of the *European Journal of Social Psychology* and has been associate editor of *Personality and Social Psychology Bulletin*.

Note

* Correspondence: Department of Social Psychology, University of Groningen, Groningen, The Netherlands. Email: s.otten@rug.nl

References

- Abrams, D., Randsley de Moura, G., Marques, J. M., & Hutchison, P. (2008). Innovation credit: When can leaders oppose their group's norms? *Journal of Personality and Social Psychology*, 95, 662–678. doi: 10.1037/0022-3514.95.3.662
- Abrams, D., Randsley de Moura, G., & Travaglino, G. A. (2013). A double standard when group members behave badly: Transgression credit to ingroup leaders. *Journal of Personality and Social Psychology*, 105, 799–815. doi: 10.1037/a0033600
- Amodio, D. M., & Mendoza, S. A. (2010). Implicit intergroup bias: Cognitive, affective, and motivational underpinnings. In B. Gawronski, & B. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications* (pp. 353–374). New York: Guilford Press.
- Baumeister, R. F. (1998). The self. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology* (Vols. 1 and 2, 4th edn., pp. 680–740). New York: McGraw-Hill.
- Baumeister, R. F., Tice, D. M., & Hutton, D. G. (1989). Self-presentational motivations and personality differences in self-esteem. *Journal of Personality*, 57, 547–579. doi: 10.1111/j.1467-6494.1989.tb02384.x
- Biernat, M., Vescio, T. K., & Billings, L. S. (1999). Black sheep and expectancy violation: Integrating two models of social judgment. *European Journal of Social Psychology*, 29, 523–542. doi: 10.1002/(SICI)1099-0992(199906)29:4<523::AID-EJSP944>3.0.CO;2-J
- Braun, M. (2010). Dealing with a deviant group member. Kurt Lewin Institute Dissertation Series, 2010–12. Enschede, NL: GildePrint.
- Braun, M., Otten, S., & Gordijn, E. H. (2009a). Did you really mean that? The role of group affiliations and ambiguity of negative intentions in aggressive interactions. *Unpublished manuscript*, University of Groningen, The Netherlands.
- Braun, M., Otten, S., & Gordijn, E. H. (2009b). The impact of cognitive load on reactions to deviant ingroup members. *Unpublished manuscript*, University of Groningen, The Netherlands.

- Brewer, M. B. (1979). Ingroup bias in the minimal intergroup situations: A cognitive motivational analysis. *Psychological Bulletin*, 86, 307–324. doi: 10.1037//0033-2909.86.2.307
- Brewer, M. B., & Brown, R. J. (1998). Intergroup relations. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vols. 1 and 2, 4th edn., pp. 554–594). New York: McGraw-Hill.
- Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin*, 135, 183–204. doi: 10.1177/0022002710383670
- Frijda, N. H. (1986). *The Emotions*. Cambridge, England: Cambridge University Press.
- Gilbert, D. T., & Hixon, J. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, 60, 509–517. doi: 10.1037/0022-3514.60.4.509
- Haslam, S. A. (2001). *Psychology in Organizations: The social Identity Approach*. London, England: Sage.
- Hewstone, M. (1990). The 'ultimate attribution error'? A review of the literature on intergroup causal attribution. *European Journal of Social Psychology*, 20, 311–335. doi: 10.1002/ejsp.2420200404
- Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, 53, 575–604. doi: 10.1146/annurev.psych.53.100901.135109
- Hogg, M. A. (2001). A social identity theory of leadership. *Personality and Social Psychology Review*, 5, 184–200. doi: 10.1207/S15327957PSPR0503_1
- Hunter, J. A., Stringer, M., & Watson, R. P. (1991). Intergroup violence and intergroup attributions. *British Journal of Social Psychology*, 30, 261–266. doi: 10.1111/j.2044-8309.1991.tb00943.x
- Karelalia, N., & Keck, S. (2013). When deviant leaders are punished more than non-leaders: The role of deviance severity. *Journal of Experimental Social Psychology*, 49, 783–796. doi: 10.1016/j.jesp.2013.04.003
- Lieberman, J. D., Solomon, S., Greenberg, J., & McGregor, H. A. (1999). A hot new way to measure aggression: Hot sauce allocation. *Aggressive Behavior*, 25, 331–348. doi: 10.1002/(SICI)1098-2337(1999)25:5<331::AID-AB2>3.0.CO;2-1
- Marques, J. M., & Yzerbyt, V. Y. (1988). The black sheep effect: Judgmental extremity towards ingroup members in inter- and intra-group situations. *European Journal of Social Psychology*, 18, 287–292. doi: 10.1002/ejsp.2420180308
- Marques, J., Abrams, D., Paez, D., & Martinez-Taboada, C. (1998). The role of categorization and in-group norms in judgments of groups and their members. *Journal of Personality and Social Psychology*, 75, 976–988. doi: 10.1037/0022-3514.75.4.976
- Marques, J., 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. doi: 10.1037/0022-3514.81.3.436
- Marques, J. M., Robalo, E. M., & Rocha, S. A. (1992). Ingroup bias and the 'black sheep' effect: Assessing the impact of social identification and perceived variability on group judgements. *European Journal of Social Psychology*, 22, 331–352. doi: 10.1002/ejsp.2420220403
- Marques, J. M., Yzerbyt, V. Y., & Leyens, J. (1988). The 'Black Sheep Effect': Extremity of judgments towards ingroup members as a function of group identification. *European Journal of Social Psychology*, 18, 1–16. doi: 10.1002/ejsp.2420180102
- Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology*, 22, 103–122. doi: 10.1002/ejsp.2420220202
- Mummendey, A., & Otten, S. (1993). Aggression: Interaction between individuals and social groups. In R. B. Felson, & J. T. Tedeschi (Eds.), *Aggression and Violence: Social Interactionist Perspectives* (pp. 145–167). Washington, DC: American Psychological Association. doi: 10.1037/10123-006
- Mummendey, A., Linneweber, V., & Löschper, G. (1984). Actor or victim of aggression: Divergent perspectives – divergent evaluations. *European Journal of Social Psychology*, 14, 297–311. doi: 10.1002/ejsp.2420140305
- Otten, S., & Gordijn, E. H. (2008). Getting away with deviant behavior: The impact of social categorization and (loss of) power. Paper presented at the General Meeting of the European Association of Social Psychology, Opatija, Croatia, June 2008.
- Otten, S., & Moskowitz, G. B. (2000). Evidence for implicit evaluative in-group bias: Affect-biased spontaneous trait inference in a minimal group paradigm. *Journal of Experimental Social Psychology*, 36, 77–89. doi: 10.1006/jesp.1999.1399
- Otten, S., & Wentura, D. (1999). About the impact of automaticity in the Minimal Group Paradigm: Evidence from affective priming tasks. *European Journal of Social Psychology*, 29, 1049–1071. doi: 10.1002/(SICI)1099-0992(199912)29:8<1049::AID-EJSP985>3.0.CO;2-Q
- Pinto, I. R., Marques, J. M., Levine, J. M., & Abrams, D. (2010). Membership status and subjective group dynamics: Who triggers the black sheep effect? *Journal of Personality and Social Psychology*, 99, 107–119. doi: 10.1037/a0018187
- van Prooijen, J. (2006). Retributive reactions to suspected offenders: The importance of social categorizations and guilt probability. *Personality and Social Psychology Bulletin*, 32, 715–726. doi: 10.1177/0146167205284964
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin, & S. Worchel (Eds.), *The Psychology of Intergroup Relations* (pp. 33–47). Monterey, CA: Brooks/Cole.