This is a repository copy of *The Whistleblower's Dilemma in Young Children: When Loyalty Trumps Other Moral Concerns*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/127610/

Version: Accepted Version

Article:

The Whistleblower's Dilemma in Young Children: When Loyalty Trumps Other Moral Concerns

Antonia Misch[a,d], Harriet Over[b], & Malinda Carpenter[a,c]

[a] Department of Developmental and Comparative Psychology,
Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
[b] Department of Psychology, University of York, York, UK
[c] School of Psychology and Neuroscience, University of St Andrews, St Andrews, UK
(present address)
[d] Department of Psychology, Ludwig Maximilian University, Munich, Germany
(present address)

Word Count: 5,485

Correspondence:
Antonia Misch
Department of Psychology
Ludwig Maximilian University
Leopoldstr. 13
80802 Munich
Germany
e-mail: antonia.misch@psy.lmu.de
phone: +49 (89) 2180 6048

Abstract

When a group engages in immoral behavior, group members face the whistleblower's dilemma: the conflict between remaining loyal to the group and standing up for other moral concerns. This study examines the developmental origins of this dilemma by investigating 5-year-olds' whistleblowing on their in- vs. outgroup members' moral transgression. Children (n=96) watched puppets representing their ingroup vs. outgroup members commit either a mild or a severe transgression. After the mild transgression, children tattled on both groups equally often. After the severe transgression, however, they were significantly less likely to blow the whistle on their ingroup than on the outgroup. These results suggest that children have a strong tendency to act on their moral concerns, but can adjust their behavior according to their group's need: When much is at stake for the ingroup (i.e., after a severe moral transgression), children's behavior is more likely to be guided by loyalty.

Keywords:
intergroup cognition, group loyalty, morality, whistleblowing, social cognition

In press at *Frontiers in Psychology*

PLEASE DO NOT DISTRIBUTE WITHOUT PERMISSION

The whistleblower's dilemma in young children: When loyalty trumps other moral concerns

During recent years, high profile cases of whistleblowing have garnered enormous public attention and caused controversy in politics and the international media. For example, recently, the former CIA contractor Edward Snowden, who revealed top-secret information about surveillance programs run by the US National Security Agency, was extensively both reviled and lauded in equal measure for being a whistleblower. Whistleblowing is the disclosure of one's own group's transgressions with the intention of stopping the group's wrongdoing, which necessarily involves an act of disloyalty against the group (see Jubb, 1999). Whistleblowers thus experience a dilemma in which they have to decide whether to act on their feelings of group loyalty or on other moral principles (Waytz, Dungan, & Young, 2013). According to Haidt and colleagues, loyalty is one of five moral foundations (Haidt & Joseph, 2007) and requires preferential treatment for members of one's own group. In contrast, other moral concerns, such as fairness and care, demand equal treatment for all (Haidt & Graham, 2007). Thus, loyalty can involve sacrificing other moral principles to protect the group, while whistleblowing involves privileging these other moral concerns over loyalty. The consequences of whistleblowing for both the group and the whistleblower can often be severe. The group may be punished externally, and the whistleblower may be punished by the group as a traitor, and maybe even excluded or banned.

Surprisingly, the conditions under which people decide whether to blow the whistle on their group have not been extensively investigated. Research with adults has examined the effects of factors such as the interests of the group and role responsibility (Trevino & Victor, 1992), or monetary incentives and legal protections (Oh & Teo, 2010). Other studies have focused on whistleblowing in interpersonal rather than group contexts (e.g., Bocchiaro, Zimbardo, & Van Lange, 2012; Gino & Bazerman, 2009; Waytz et al., 2013). Only a small amount of research has directly investigated the effects of morality and loyalty concerns on whistleblowing. A set of studies conducted by Waytz and colleagues (2013) suggests that participants' willingness to blow the whistle on another person is predicted by their endorsement of fairness over loyalty concerns. They also found that participants' willingness to blow the whistle decreases with closeness between the participant and the transgressor. It is not yet clear, however, what happens when loyalty and other moral concerns are directly pitted against each other in a group context. Furthermore, a common feature of previous research is that it has assessed participants' predictions of how they might act if faced with this dilemma. But evidence suggests that participants' predictions can diverge from their actual behavior. For example, in a study conducted by Bocchiaro and colleagues (2012), a large majority of participants predicted that they would blow the whistle on an unethical request, but only a small minority actually did so when put to the test, stressing the importance of investigating the whistleblowing dilemma in a behavioral set-up.

Developmental research has shown that both components of the whistleblower's dilemma, feelings of group loyalty and other moral concerns, are present early in childhood. At least by 5 years of age, children clearly value loyalty to the group: They favor loyal over disloyal group members (e.g., Abrams, Rutland, & Cameron,

2003; Abrams, Rutland, Ferrell, & Pelletier, 2008; Misch, Over, & Carpenter, 2014). They also show loyal behavior themselves, even when it is costly for them to do so (Misch, Over, & Carpenter, 2016). Young children are also sensitive to other basic moral principles. For example, from the age of 3 years, children actively intervene in moral transgressions in which a third party has been harmed (Rossano, Rakoczy, & Tomasello, 2011; Vaish, Missana, & Tomasello, 2011), give more resources to an individual who behaved in a morally good way (Kenward & Dahl, 2011), and avoid helping people with harmful intentions (Vaish, Carpenter, & Tomasello, 2010). They are also concerned with fairness, for example they prefer a fair to an unfair distributor in a third party context (e.g., Shaw, DeScioli, & Olson, 2012).

However, to our knowledge, no study has directly investigated the conflict between loyalty and other moral considerations in an intergroup context in young children. The few studies that have investigated the related issue of the interplay between ingroup favoritism[1] and fairness in children have found mixed results. DeJesus, Rhodes, and Kinzler (2014) found that in a third-party context, at least from 6 years of age, children expect others to favor their ingroup, but evaluate fair distributions as nicer. However, when evaluating their own ingroup members' resource distributions between groups, Cooley and Killen (2015) found that 3.5- to 6-year-old children value fairness over group considerations, whereas Jordan, McAuliffe, and Warneken (2014) found that 6- and 8-year-old children tended to decide whether to punish unfair distributors based on group membership, and in doing so, sacrificed moral considerations that demand equal treatment for all (Rhodes & Chalik, 2013).

The studies that have come closest to investigating the conflict between loyalty and other moral concerns are studies on so-called "blue lies" – the opposite of whistleblowing – that is, lies that are told to protect someone else. Several studies have investigated children's evaluations of blue lies in story vignettes and found that with age, children evaluate blue lies to cover up the ingroup's transgression more positively (e.g., Chiu Loke, Heyman, Itakura, Toriyama, & Lee, 2014; Fu, Luo, Heyman, Wang, Cameron, Lee, 2016; Lau, Cameron, Chieh, O'Leary, Fu, & Lee, 2013; Sweet, Heyman, Fu, & Lee, 2010). To our knowledge, only one behavioral study has directly focused on children's blue lies by asking participants to report their own group's wrongdoing. Fu, Evans, Wang, and Lee (2008) tempted class groups of 7- to 11-year-old Chinese children to cheat in a competition by allocating more expert players to their team than were allowed. Afterwards, an uninvolved experimenter asked children in a confidential one-to-one situation whether their team really played by the rules. The majority of children confessed their team's transgression and thus acted according to their moral considerations rather than their feelings of loyalty. However, this study did not include an outgroup comparison so it is not known whether children would have been even less likely to lie for an outgroup. It thus still remains open how children would weigh moral and loyalty concerns when deciding what to do about an ingroup versus an outgroup member's transgression.

---

[1] Note that loyalty to the group is more than simple ingroup preference, in that it entails a sense of commitment and often the willingness to sacrifice personal benefits for the sake of the group (see Brewer & Silver, 2000, p. 162).
[2] For 21 of these excluded children it was still possible to obtain a tattling score. Adding them to the

A promising approach to study this conflict is to look at children's tattling behavior. The terms tattling and whistleblowing are often used interchangeably (see e.g., Waytz et al., 2013), but one important distinction will be made here: While tattling can be used rather generally and independently of group membership or affiliations (see e.g., Ingram & Bering, 2010), whistleblowing refers specifically to tattling about one's own organization or group (e.g., Jubb, 1999). For children, tattling is a frequent and natural way of dealing with others' transgressions and misbehavior. Young children do not perceive tattling as negative and thus frequently tattle on peers in school (Ingram & Bering, 2010), on their siblings (Den Bak & Ross, 1996), on puppets in experimental settings (Schmidt, Rakoczy, & Tomasello, 2012; Vaish et al., 2011), and even on adults' transgressions (Heyman, Chiu Loke, & Lee, 2016).

To investigate the origins of the whistleblower's dilemma in young children, we thus study children's tattling behavior. Children observed either ingroup or outgroup members commit a moral transgression. Afterwards, an uninvolved experimenter entered the room and gave children the chance to spontaneously tattle before asking more direct questions. We expected that children would be more likely to tattle on the outgroup's than on the ingroup's transgression, because previous developmental research has shown that young children are loyal to their groups (Misch et al., 2016) and research with adults has shown that the closeness of one's relationship to the transgressor is negatively correlated with the likeliness to blow the whistle on him/her (Waytz et al., 2013). We chose to test 5-year-old children because this is the earliest age at which clear evidence exists that children both value loyalty to the group (Misch et al., 2014, 2016) and are concerned about moral transgressions (Blake & Harris, 2009; Rossano et al., 2011; Vaish et al., 2011).

Additionally, to investigate the conflict between loyalty and morality more deeply, we were also interested in the impact of the severity of the transgression. More specifically, we wished to examine whether and how loyalty and the severity of the moral violations would interact. Results from previous studies looking at children's evaluations of and reactions to different types of transgressions are mixed. One line of research has found that 4- to 7-year-old children endorse tattling on both major and minor transgressions equally, and only from around 8-9 years endorse tattling on major more than on minor transgressions (Chiu Loke, Heyman, Forgie, McCarthy, & Lee, 2011; Chiu Loke et al., 2014; Heyman, Chiu Loke, & Lee, 2016; Lyon, Ahern, Malloy, & Quas, 2010). However, in these studies children were simply asked to evaluate or predict vignette story characters' tattling behavior. Behavioral studies that have investigated children's own behavior following different types of transgressions have found that by 3 years of age, children differentiate between severe moral transgressions and more minor conventional violations in that they protest more strongly when someone destroys the possession of another person compared to when someone plays a game incorrectly (Schmidt et al., 2012).

In the current study, the transgression was implemented in the form of a theft. Previous research has shown that from around 3 years of age, children understand the violation of property rights and protest against this (Rossano et al., 2011). At least by age 5, they understand the illegitimate nature of stealing (Blake & Harris, 2009). An advantage of using this type of transgression is that it allowed for a quantitative manipulation of severity: Children in the mild transgression condition observed two

puppets take only a little bit of someone's possession (i.e., 1 out of 10 gemstones), while children in the severe transgression condition observed these puppets take nearly all of that resource (i.e., 9 out of 10 gemstones). For children for whom these two puppets were outgroup members (outgroup condition) we expected generally high levels of tattling in both transgression conditions (although they might tattle more in the severe transgression condition). Children should not feel any loyalty to the outgroup members, and therefore should act according to their moral considerations and, consequently, tattle. For children for whom the transgressors were ingroup members (ingroup condition), observing the mild versus severe transgression should also elicit mild versus severe moral considerations; however in this case these considerations should conflict with loyalty considerations. We expected that children's feelings of group loyalty would make it more difficult for them to blow the whistle on their ingroup members. There were two different possible ways in which the severity of the transgression might influence their behavior.

The first possibility was that in the mild condition, compared to their feelings of loyalty, children's moral considerations should be relatively low, and consequently children might act according to their feelings of loyalty and keep quiet about their group's transgression. After the severe transgression, however, moral considerations should outweigh feelings of loyalty, and thus children might act on their moral considerations and blow the whistle on their group.

The social psychological literature with adults suggests a second possibility. According to a nonabandonment norm, group members should stick to their group in all circumstances (Zdaniuk & Levine, 2001), but especially in situations in which it is needed most (e.g., because the group is under threat; see Ellemers, Spears, & Doosje, 2002; Van Vugt & Hart, 2004). Indeed, some evidence supports the notion that threat to the group increases group cohesion or ingroup bias (Hunter et al., 2005; Turner, Hogg, Turner, & Smith, 1984), and that after undergoing negative experiences, group members feel more fused with each other (e.g., Jong, Whitehouse, Kavanagh, & Lane, 2015) and show more pro-group behavior (Swann et al., 2014). If this is the case, then children should keep quiet after their own group's severe transgression, as otherwise the group members would have to face punishment or other negative consequences. After a mild transgression, in contrast, potential negative consequences should be relatively minor and not harm the ingroup much; therefore children could act according to their moral considerations and blow the whistle.

Method

Participants

Participants were 96 five-year-old children (48 girls and 48 boys, age range 5;0;27 - 5;9;09, $M = 5$ years, 6 months). The number of participants (24 per condition) was specified in advance based on previous research (Misch et al., 2016). Twenty-two additional children were tested but excluded for failing one of the critical control questions that tested whether they understood the procedure (i.e., failing to correctly say which group they were in [1], failing to correctly say which group the transgressors were in [4], or failing to remember whether one vs. many gemstones were taken away [8]), or for experimenter error (5), not responding at all (1), leaving

the room during the procedure (2), or naming one of the transgressors after herself (1).[2]

Children were recruited and tested in their daycare centers in a mid-sized city in Germany. The test session took approximately 20 minutes. No SES or ethnicity data were collected, but approximately 98% of the population from which the sample was drawn are native German. The study was developed and conducted in accordance with ethical guidelines and was approved by the institution's ethics committee (Max Planck Institute for Evolutionary Anthropology Child Subjects Committee).

Materials

We used puppets as in- and outgroup members because previous work has shown that children are willing to tattle on puppets' transgressions (e.g., Vaish et al., 2011). Children were tested by three female experimenters: a moderator (M) and two puppeteers (E1 and E2). Each puppeteer played one female and one male hand puppet (see Figure 1). The two puppets played by E1 were the transgressors. In the ingroup condition the child was allocated to the same group as the transgressors; in the outgroup condition the child was allocated to the other group.

A set of green and yellow scarves (two puppet-sized scarves and a child-sized scarf in each color; see Figure 1) were stored in a box with a lid. Ten fake red gemstones were used as spoils (see Figure 2). They were hidden in a small purse located on a box on the left side of the room (approximately 2m away from the door).

Figure 1. The puppets used in this study: a) The transgressors played by E1 (here wearing green group markers), b) The puppets played by E2 (here wearing yellow group markers).

There was a low cardboard barrier (30cm in height) on the other side of the room. Thirty large marbles and a marble bag were used to keep children occupied and in place before and during the transgression, and a marble run was used for the preference test at the end.

---

[2] For 21 of these excluded children it was still possible to obtain a tattling score. Adding them to the main analysis did not change the results (full-null model comparison: $\chi^2(3) = 8.04$, $p = .045$; interaction between group membership and transgression type: *Estimate* = 2.84, *SE* = 1.14, $\chi^2(1) = 7.08$, $p < .01$).

Figure 2. The ten gemstones used as spoils.

Design and Counterbalancing

Children were tested in a 2x2 between-subjects design. We manipulated the transgressors' group membership and the transgression type: Transgressors were either in the child's in- or outgroup, and took either a little (only one out of 10 gemstones in the mild condition) or a lot (nine out of 10 gemstones in the severe condition).

Across children, we counterbalanced the color of the child's group (so that half of the children in each condition were in the yellow group, and the other half were in the green group) and the color of the transgressors' group (so that half of the time they were in the yellow group, and half of the time they were in the green group).

Procedure

Children were picked up from their classroom individually by all three experimenters. At the start of the procedure, there was a brief warm-up phase during which children became acquainted with the experimenters and the four puppets that would later be allocated to groups. First, the moderator (M) introduced the child to the puppets and then asked the puppets to introduce themselves. Following this, M asked the child and each of the puppets two questions to engage them in a brief conversation (e.g., about what they had had for breakfast, or which parent dropped them off at the daycare). This was done in order to make the child feel comfortable in the situation and to establish that the puppets should be treated as if they were real individuals around the same age as the child.

Group allocation. After the warm-up, M allocated the child and the four puppets to groups. She did this by saying, "Today, we need two different groups. We will have a yellow group and a green group. First of all, we need to know which group everyone belongs to." M then picked up the box and explained that in this box there were yellow and green scarves, and that she would now pull out one scarf for each of them, thereby finding out which group they belonged to. Then, one by one, she allocated each of the puppets and the child into groups by apparently randomly

drawing yellow and green scarves out of the box and placing them on each individual's neck. Group allocation always started with one of the child's ingroup puppets, then proceeded to an outgroup puppet, then to the child, the other outgroup puppet, and finally the other ingroup puppet.

Transgression. After the group allocation, M said that next they would need the marbles that were lying on the floor behind the low barrier in one corner of the room. She noticed that the marble bag was missing and asked the child to come with her to look for the bag outside of the room. This was an excuse so that E1 and E2 could leave the room unseen and wait in an adjacent room. When M and the child returned with the marble bag, M pretended to be surprised that the others were missing and asked the child to put all the marbles into the bag while she looked for the others outside. The task of putting the marbles into the bag was given to children so that they would be occupied with a simple activity on one side of the room, but would still be attentive enough to observe the transgression. While the child was busy picking up the marbles, the two puppets played by E1 entered the room. Depending on condition, they were either in the same group as the child (ingroup condition) or in the other group (outgroup condition). They recognized and greeted the child very briefly, before turning to each other and ignoring the child. The male puppet then said, "Look, there is the purse! Maybe there are gemstones in it again, and we could take some again!" To make sure that children understood that the puppets were not entitled to take the gemstones, the other puppet was skeptical and pointed out that the gemstones did not belong to them. In order to convey the idea that this was something this group did regularly, the first puppet said "But we are members of the yellow/green group, and the yellow/green group always does it like that!" The female puppet then replied, "Ok, then let's have a look. But let's be quick and quiet, so that no one will catch us!" They then opened the purse and admired the gemstones. Depending on the condition, they took either one (mild transgression) or nine of the ten gemstones (severe transgression). In the mild condition they said to each other, "Let's take only a little bit, only one gemstone. There are still many left, certainly no one will notice!" In the severe condition they said, "Let's take a lot of them, nearly all the gemstones. There is still one left, certainly no one will notice!" After they put the gemstone(s) into their purse, M called them from the outside, "[Transgressor puppets' names], where are you?" The puppets replied to M, "We are coming," and then said to themselves, "Let's leave quickly, so that no one will catch us!" Finally, before leaving the room, they asked the child to wait inside.

Tattling opportunity. Then, M entered the room and gave the child the chance to tattle. In order to assess how quickly and spontaneously children tattled, she used a stepwise, ramping-up procedure with a 5-second pause in between each step to give children time to tattle. She first started with very general comments (Step 0: "I'm back" and Step 1: "Is everything okay?"), and then gave some hints that something was amiss while looking at the bag (Step 2: "What is going on here?" and Step 3: "What did I miss?"). She then asked more directly about the bag (Step 4: "There is a bag. Someone must have forgotten it..." and Step 5: "The zipper is open. Maybe someone took something out?"). In the final step she finally suspected the puppets directly (Step 6: "I think I just saw [names of transgressor puppets] leave, maybe they took something?"). If the child did not respond at all during a given 5-second response period, or only said something unrelated (e.g., just talked about the marbles that they had picked up), M moved on to the next step. For children's

statements to qualify as tattling, they had to make it clear that someone had taken something away (for step 6 it was sufficient if they confirmed M's suspicion by saying "yes"). If children only gave a hint of this, M further encouraged them by saying "Uh huh, tell me!" If children correctly described what had happened but failed to name the transgressor(s), M asked "And who?" Following that, children had another 5 seconds until, if needed, M moved on to the next step. To minimize social pressure on children, M looked only briefly at them and then continued to inspect the scene. Thus, children were free to remain silent.

Post-test measures and resolution. To explore the motivation underlying children's behavior, we asked them some post-test questions about their justification for and evaluation of the transgression, their judgment of the transgressors, their own accountability (only for tattlers), their loyalty, and their group preference. Because these questions were exploratory, we did not push children to answer if they did not respond. As a consequence the number of *no answer* responses was relatively high and the results should be taken with some caution. Furthermore, grouping children depending on their tattling behavior led to small and uneven sample sizes in the different cells. Thus, for most of the measures, statistical analysis was not appropriate; therefore we report these results in the supplementary material.

Memory questions. After the first set of post-test questions (but before the loyalty question), in order to make sure that children had followed and understood the procedure, M asked children three memory questions: "Which group did the two who took the stones belong to?", "How many stones did they take, many or just a few?", and "Which group do you belong to?"

Resolution of the situation. After M's loyalty question, before the preference test, the two transgressor puppets re-entered the room. They were clearly upset by their wrongdoing, confessed everything to M, and apologized. M explained that taking away others' belongings is not okay and made them promise never to do anything like this again. Then the other two puppets came back and everyone played together with a marble run. Finally, children were thanked and taken back to their classroom.

Coding and Reliability

Our main interest was in whether children tattled on the puppets' transgression or not (saying, e.g., "They took the gemstones" or "I saw that they stole something"). Children's statements were coded as tattling if they made it clear that someone (e.g., "they", "the puppets", "the two", or using their names) had taken something away (e.g., "They took something", "They swiped the stones"). Only for Step 6 was it sufficient if they clearly confirmed M's suspicion by saying "yes").

In addition, for those children who tattled, we also investigated how quickly and spontaneously they tattled. For this analysis, children received a score between 0 and 7, corresponding to the step at which they tattled (e.g., they received 1 if they tattled at step 1, or 6 if they tattled at step 6). If they tattled before M's first hint, they received a 0, and if they did not tattle at all, they received a score of 7.

The main coding was done by the first author. To assess inter-rater reliability, an independent coder who was unaware of the hypotheses of the study coded a

random sample of 25% of children for both measures together from the videos. Reliability (Cohen's weighted kappa) was perfect with κ = 1.00.

Results

All statistical analyses were performed using R (R Core Team, 2014) version 3.2.0. Significance of the models was tested using both likelihood ratio tests (LRT), by comparing the fit of the full model with that of the respective reduced models, and the p values provided by the final model.

A preliminary analysis revealed no effects of children's gender or color group on the main results regarding children's tattling (General Linear Model, full-null model comparison, $p > .25$). Therefore, we collapsed across these variables and do not consider them further.

Our main interest was in how many children tattled about the puppets' transgression at any point during the test phase. Overall, across all four conditions, the majority of children tattled (82.3%), suggesting a general concern for harm. Figure 3 depicts the percentage of children who tattled in each condition for each transgression type.
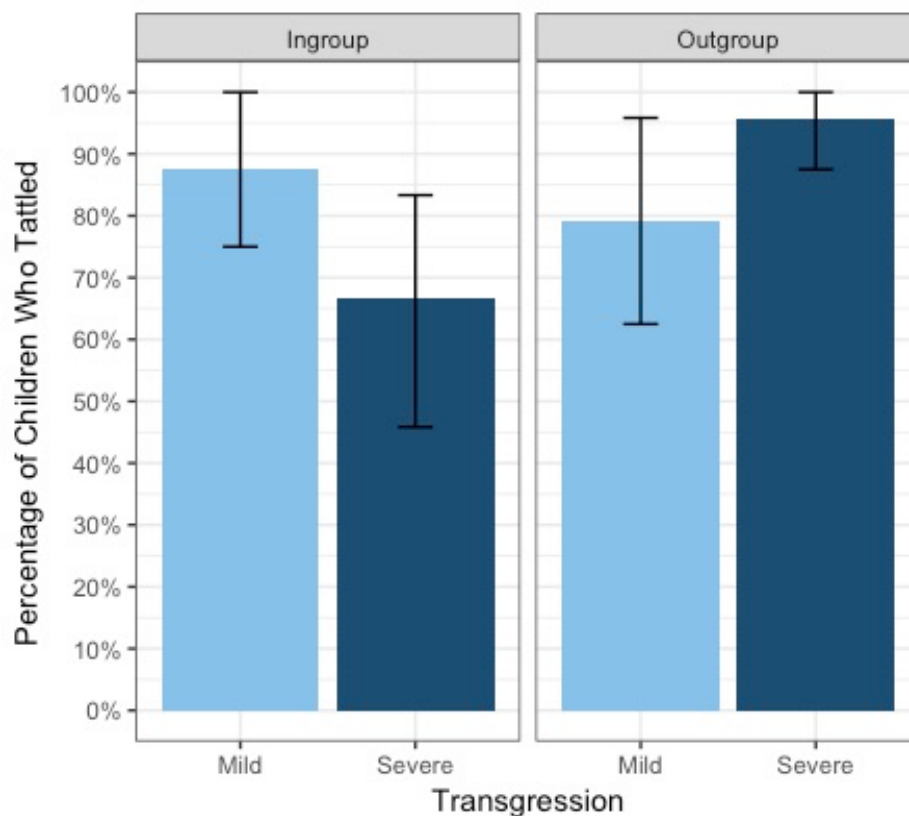

Figure 3. Percentage of children who tattled, by group membership and transgression type, with bootstrapped 95% confidence intervals.

A GLM was run with group membership and transgression type as predictors, and the binomial measure of tattling (yes or no) as response variable. The full model differed significantly from the null model ($\chi^2(3) = 8.14$, $p = .043$) and revealed a

significant interaction between group membership and transgression type (*Estimate* = 3.05, *SE* = 1.36, $\chi^2(1) = 6.26$, *p* = .012, Nagelkerke's $R^2 = .11$). Post-hoc tests revealed that children in the ingroup condition were significantly less likely to tattle on a severe transgression than were children in the outgroup condition (Fisher's exact test, *p* = .023, risk ratio = 2.17); all other pairwise comparisons were non-significant (Fisher's exact tests, all *p*'s > .16).

To investigate whether the conditions had an effect on how quickly children tattled, we ran a GLM with Poisson distribution only on children who had tattled at some point (*n* = 79), with children's tattling score (0 - 6) as the response variable. The full model did not differ from the null model (*p* = .21), indicating that the conditions had no significant effect on how quickly children tattled (see Figure 4).
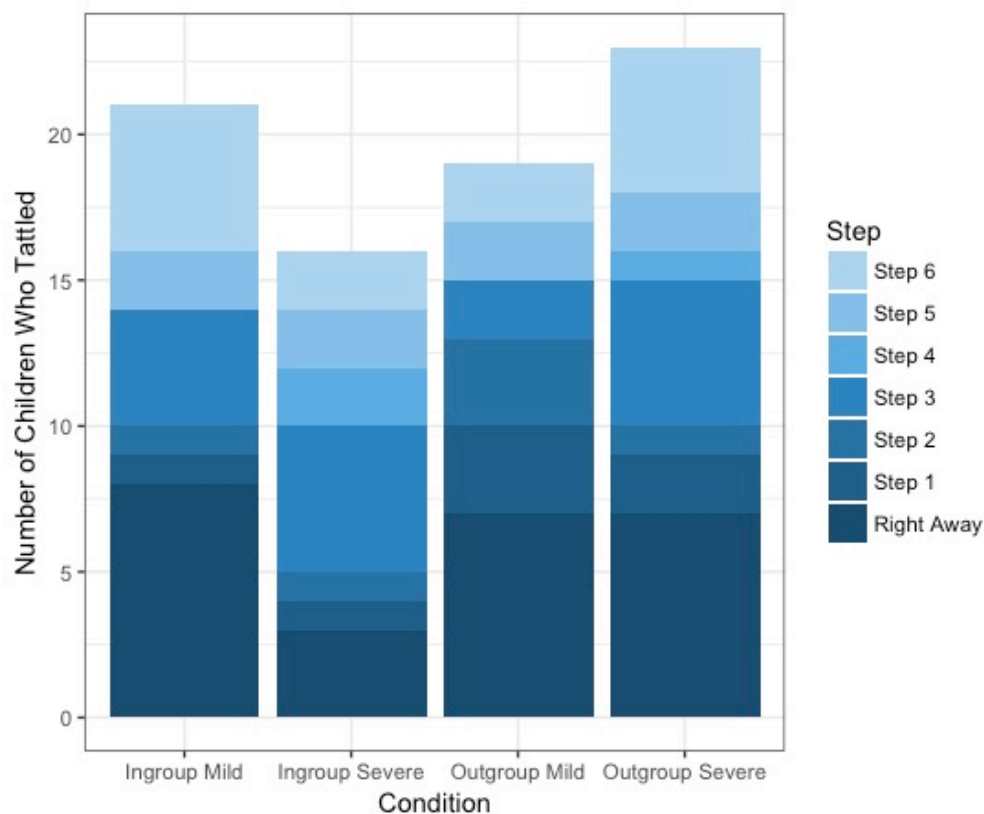

Figure 4. Number of children who tattled at each step, by condition.

Discussion

The aim of this study was to examine the whistleblower's dilemma: the conflict between feelings of loyalty and other moral concerns. This was done by looking at children's willingness to blow the whistle on their in- versus outgroup members' mild versus severe transgression. An interesting pattern of results emerged: Rather than simply tattling more on outgroup members across the board, children showed a complex weighting of loyalty and moral considerations. After the mild transgression, children acted on their moral considerations: They tattled on both groups at equally high rates. After the severe transgression, however, they were significantly less likely to blow the whistle on their ingroup than on the outgroup, suggesting that children's feelings of loyalty to the group sometimes outweighed

other moral considerations. Consistent with the idea of a nonabandonment norm, these results support the notion that group loyalty becomes most important when much is at stake for the group, that is, that one should show the strongest loyalty when the group is under threat (Ellemers et al., 2002; Van Vugt & Hart, 2004). These results therefore suggest that young children are already capable of flexibly weighing moral and loyalty considerations and, in some cases, are willing to sacrifice their moral principles for the sake of their group.

There are a number of possible motivations that could have been underlying children's loyal behavior. We aimed to investigate these with the questions we asked children following the tattling phase.   Unfortunately, these findings were underpowered due to low response rates (see supplementary materials). Still, some of our and others' findings can shed light on the possible underlying motivations. For example, it is possible that children generally perceive transgressions of their ingroup members as less severe than transgressions of their outgroup members and that this led them to blow the whistle less often in the ingroup condition. Previous research has shown that children are more forgiving and forgetful when it comes to the negative behavior of their ingroup members (Corenblum, 2003; Dunham et al., 2011). However, in the mild condition of the current study children were equally likely to blow the whistle on their in- and outgroup members, suggesting that this factor alone cannot explain the observed results. Another possibility is that children might have wanted to protect their ingroup from the potential negative consequences (e.g., punishment) of their whistleblowing and/or avoid being punished themselves. Previous work has shown that children feel responsible for their group members' negative actions (Over, Vaish, & Tomasello, 2016), and consequently children's feelings of shame or embarrassment about their group members' transgression might have decreased their whistleblowing in the ingroup severe condition. Relatedly, some children might have been shocked about their group's transgression and thus too preoccupied to speak out about it. Future research should thus investigate the role of moral emotions such as guilt and embarrassment, and also the fear of negative consequences in the context of loyal behavior. Another potential reason for children's increased loyalty after the severe transgression is the fact that the group was now in possession of the stolen goods. Previous research has shown that children prefer wealthy over less wealthy groups (Horwitz, Shutts, & Olson, 2014) and show more loyalty to groups that are of high status (e.g., Nesdale & Flesser, 2001). However, children's justifications for why they wanted to stay in or leave their group suggest that this was not the reason for their choice in this situation: No child ever justified their choice to stay with or join the transgressors' group by mentioning the group's wealth, higher status, or the possession of the stones more generally, while the transgression was a common reason for joining or not leaving the non-transgressors' group.

In future research, it would be informative to investigate in which situations children are willing to override their moral concerns in order to remain loyal. It would be interesting to look at children's loyalty after even more different types of transgressions, including a wider range of severity and different kinds of moral violations. Also, if children were asked to choose between internal within-group protest (i.e., scolding and correcting ingroup members privately) and external tattling (i.e., telling someone outside the group more publicly), would their choice depend on their group membership?

In summary, our findings suggest that both loyalty and other moral considerations guide 5-year-old children's behavior. When moral concerns are relatively low, children act freely on them by tattling on the outgroup and even blowing the whistle on their own group. In contrast, when moral concerns increase, children's behavior is guided by their loyalty: They tattle freely on their outgroup, but are less likely to blow the whistle on their own group. Thus, already by 5 years of age, children consider both loyalty and other moral concerns together, and adapt their behavior flexibly. Even though they clearly understood the negative nature of the transgression, they were willing to sacrifice their personal moral concerns for the sake of their group. This is an interesting finding, given the fact that from very early on, children show a strong appreciation for key moral domains such as care and fairness (e.g., Hamlin & Wynn, 2011; Hamlin, Wynn, & Bloom, 2007; Vaish et al., 2011), while robust preferences for minimal ingroups and clear loyal behavior do not appear much before the age of five (Dunham et al., 2011; Misch et al., 2016). Thus, right around the time that loyalty to the group first appears in ontogeny, it can already have a dark side, overriding other moral concerns. This can lead to rather undesirable behavior on the one hand, for example when it results in concealing moral transgressions of the ingroup. However, from the perspective of the group, it may be seen as desirable in that it helps ensure from early on that group members are trustworthy and protective of their group and thus that they can be counted on when most needed.

References

Abrams, D., Rutland, A., & Cameron, L. (2003). The development of subjective group dynamics: Children's judgments of normative and deviant in-group and out-group individuals. *Child Development, 74*(6), 1840-1856. doi: 10.1046/j.1467-8624.2003.00641.x14669899

Abrams, D., Rutland, A., Ferrell, J. M., & Pelletier, J. (2008). Children's judgments of disloyal and immoral peer behavior: Subjective group dynamics in minimal intergroup contexts. *Child Development*, *79*(2), 444–461. doi: 10.1111/j.1467-8624.2007.01135.x

Blake, P. R., & Harris, P. L. (2009). Children's understanding of ownership transfers. *Cognitive Development*, *24*(2), 133-145. doi: 10.1016/j.cogdev.2009.01.002

Bocchiaro, P., Zimbardo, P. G., & Van Lange, P. A. (2012). To defy or not to defy: An experimental study of the dynamics of disobedience and whistle-blowing. *Social Influence, 7*(1), 35-50. doi: 10.1080/15534510.2011.648421

Cooley, S., & Killen, M. (2015). Children's evaluations of resource allocation in the context of group norms. *Developmental psychology, 51*(4), 554. doi: 10.1037/a0038796

Chiu Loke, I., Heyman, G. D., Forgie, J., McCarthy, A., & Lee, K. (2011). Children's moral evaluations of reporting the transgressions of peers: Age differences in evaluations of tattling. *Developmental Psychology*, *47*(6), 1757–1762. doi: 10.1037/a0025357

Chiu Loke, I., Heyman, G. D., Itakura, S., Toriyama, R., & Lee, K. (2014). Japanese and American children's moral evaluations of reporting on transgressions. *Developmental Psychology*, *50*(5), 1520–1531. doi: 10.1037/a0035993

DeJesus, J. M., Rhodes, M., & Kinzler, K. D. (2014). Evaluations versus expectations: Children's divergent beliefs about resource distribution. *Cognitive Science*, *38*(1), 178–193. doi: 10.1111/cogs.12093

Den Bak, I. M., & Ross, H. S. (1996). I'm telling! The content, context, and consequences of children's tattling on their siblings. *Social Development*, *5*(3), 292–309. doi: 10.1111/j.1467-9507.1996.tb00087.x

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children. *Child Development*, *82*(3), 793–811. doi: 10.1111/j.1467- 8624.2011.01577.x

Ellemers, N., Spears, R., & Doosje, B. (2002). Self and social identity. *Annual Review of Psychology*, *53*(1), 161–186. doi: 10.1146/annurev.psych.53.100901.135228

Fu, G., Evans, A. D., Wang, L., & Lee, K. (2008). Lying in the name of the collective good: A developmental study. *Developmental Science*, *11*(4), 495–503. doi: 10.1111/j.1467-7687.2008.00695.x

Fu, G., Luo, Y. C., Heyman, G. D., Wang, B., Cameron, C. A., & Lee, K. (2016). Moral Evaluations of Lying for One's Own Group. *Infant and Child Development*, *25*(5), 355-370. doi: 10.1002/icd.1941

Gino, F., & Bazerman, M. H. (2009). When misconduct goes unnoticed: The acceptability of gradual erosion in others' unethical behavior. *Journal of Experimental Social Psychology, 45*(4), 708-719. doi: 10.1016/j.jesp.2009.03.013

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, *20*(1), 98–116. doi: 10.1007/s11211-007-0034-z

Haidt, J., & Joseph, C. (2007). The moral mind: How five sets of innate intuitions

guide the development of many culture-specific virtues, and perhaps even modules. *The innate mind*, *3*, 367-391.

Hamlin, J. K., & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development*, *26*(1), 30–39. doi: 10.1016/j.cogdev.2010.09.001

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*(7169), 557–559. doi: 0.1038/nature06288

Heyman, G. D., Loke, I. C., & Lee, K. (2016). Children spontaneously police adults' transgressions. *Journal of Experimental Child Psychology*, *150*, 155-164. doi: 10.1016/j.jecp.2016.05.012

Horwitz, S. R., Shutts, K., & Olson, K. R. (2014). Social class differences produce social group preferences. *Developmental Science*, *17*(6), 991–1002. doi: 111/desc.12181

Hunter, J. A., Cox, S. L., O'Brien, K., Stringer, M., Boyes, M., Banks, M., Hayhurst, J. G. and Crawford, M. (2005), Threats to group value, domain-specific self-esteem and intergroup discrimination amongst minimal and national groups. *British Journal of Social Psychology, 44*, 329–353. doi: 10.1348/014466604X17939

Ingram, G. P., & Bering, J. M. (2010). Children's tattling: The reporting of everyday norm violations in preschool settings. *Child Development*, *81*(3), 945–957. doi: 10.1111/j.1467-8624.2010.01444.x

Jong, J., Whitehouse, H., Kavanagh, C., & Lane, J. (2015). Shared negative experiences lead to identity fusion via personal reflection. *PLOS ONE*, 10(12), e0145611. doi: 10.1371/journal.pone.0145611

Jordan, J. J., McAuliffe, K., & Warneken, F. (2014). Development of in-group favoritism in children's third-party punishment of selfishness. *Proceedings of the National Academy of Sciences*, *111*(35), 12710–12715. doi: 10.1073/pnas.1402280111

Jubb, P. B. (1999). Whistleblowing: A restrictive definition and interpretation. *Journal of Business Ethics*, *21*(1), 77–94. doi: 10.1023/A:1005922701763

Kenward, B., & Dahl, M. (2011). Preschoolers distribute scarce resources according to the moral valence of recipients' previous actions. *Developmental Psychology*, *47*(4), 1054–1064. doi: 10.1037/a0023869

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences*, *104*(30), 12577–12580. doi: 10.1073/pnas.0705345104

Lau, Y. L., Cameron, C. A., Chieh, K. M., O'Leary, J., Fu, G., & Lee, K. (2013). Cultural differences in moral justifications enhance understanding of Chinese and Canadian children's moral decisions. *Journal of Cross-Cultural Psychology*, *44*(3), 461-477. doi: 10.1177/0022022112453315

Lieberman, D., & Linke, L. (2007). The effect of social category on third party punishment. *Evolutionary Psychology, 5*(2), 1474-7049. doi: 10.1177/147470490700500203

Lyon, T. D., Ahern, E. C., Malloy, L. C., & Quas, J. A. (2010). Children's reasoning about disclosing adult transgressions: Effects of maltreatment, child age, and adult identity. *Child Development*, *81*(6), 1714–1728. doi: 10.1111/j.1467-8624.2010.01505.x

Misch, A. (2015). *Young Children's Loyalty to the Group* (Doctoral Dissertation, Univ. Leipzig).

Misch, A., Over, H., & Carpenter, M. (2014). Stick with your group: Young

children's attitudes about group loyalty. *Journal of Experimental Child Psychology*, *126*, 19-36. doi: 10.1016/j.jecp.2014.02.008

Misch, A., Over, H., & Carpenter, M. (2016). I won't tell: Young children show loyalty to their group by keeping group secrets. *Journal of Experimental Child Psychology, 142,* 96-106. doi: 10.1016/j.jecp.2015.09.016

Nesdale, D., & Flesser, D. (2001). Social identity and the development of children's group attitudes. *Child Development*, *72*(2), 506–517. doi: 10.1111/1467-8624.00293

Oh,. L. & Teo, H. (2010). To blow or not to blow: An experimental study on the intention to whistleblow on software piracy. *Journal of Organizational Computing and Electronic Commerce, 20*(4), 347-369. doi: 10.1080/10919392.2010.516639

Over, H., Vaish, A., & Tomasello, M. (2016). Do young children accept responsibility for the negative actions of ingroup members? *Cognitive Development*, *40*, 24-32. doi: 10.1016/j.cogdev.2016.08.004

R Core Team. (2014). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. URL: http://www.R-project.org/.

Rhodes, M., & Chalik, L. (2013). Social categories as markers of intrinsic interpersonal obligations. *Psychological Science*, *24*(6), 999–1006. doi: 10.1177/0956797612466267

Rossano, F., Rakoczy, H., & Tomasello, M. (2011). Young children's understanding of violations of property rights. *Cognition*, *121*(2), 219–227. doi: 10.1016/j.cognition.2011.06.007

Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition*, *124*(3), 325–333. doi: 10.1016/j.cognition.2012.06.004

Shaw, A., DeScioli, P., & Olson, K. R. (2012). Fairness versus favoritism in children. *Evolution and Human Behavior*, *33*(6), 736–745. doi: 10.1016/j.evolhumbehav.2012.06.001

Sweet, M. A., Heyman, G. D., Fu, G., & Lee, K. (2010). Are there limits to collectivism? Culture and children's reasoning about lying to conceal a group transgression. *Infant and Child Development*, *19*(4), 422-442. doi: 10.1002/icd.669

Swann, W. B., Gómez, Á., Buhrmester, M. D., López-Rodríguez, L., Jiménez, J., & Vázquez, A. (2014). Contemplating the ultimate sacrifice: Identity fusion channels pro-group affect, cognition, and moral decision making. *Journal of Personality and Social Psychology*, 106(5), 713–727. doi: 10.1037/a0035809

Trevino, L. K., & Victor, B. (1992). Peer reporting of unethical behavior: A social context perspective. *Academy of Management Journal*, *35*(1), 38–64. doi: 10.2307/256472

Turner, J. C., Hogg, M. A., Turner, P. J., & Smith, P. M. (1984). Failure and defeat as determinants of group cohesiveness. *British Journal of Social Psychology, 23*(2), 97-111. doi: 10.1111/j.2044-8309.1984.tb00619.x

Vaish, A., Carpenter, M., & Tomasello, M. (2010). Young children selectively avoid helping people with harmful intentions. *Child Development*, *81*(6), 1661–1669. doi: 10.1111/j.1467-8624.2010.01500.x

Vaish, A., Missana, M., & Tomasello, M. (2011). Three-year-old children intervene in third-party moral transgressions. *British Journal of Developmental Psychology*, *29*(1), 124–130. doi: 10.1348/026151010X532888

Van Vugt, M., & Hart, C. M. (2004). Social identity as social glue: The origins of group loyalty. *Journal of Personality and Social Psychology*, *86*(4), 585. doi: 10.1037/0022-3514.86.4.585

Waytz, A., Dungan, J., & Young, L. (2013). The whistleblower's dilemma and the fairness-loyalty tradeoff. *Journal of Experimental Social Psychology*, *49*(6), 1027–1033. doi: 10.1016/j.jesp.2013.07.002

Zdaniuk, B., & Levine, J. M. (2001). Group loyalty: Impact of members' identification and contributions. *Journal of Experimental Social Psychology*, *37*(6), 502–509. doi: 10.1006/jesp.2000.1474

Supplementary Materials

Procedure

To help understand the motivations behind children's responses, we asked children some exploratory questions after the main test phase. To ensure that all children felt comfortable talking about the transgression, at this point M confirmed her suspicion that the two transgressors took something, expressed at Step 6, by saying aloud to herself, "That is what must have happened."

Transgression justification. To see whether children would come up with excuses for the transgressors and, if so, to check whether the excuses would differ between conditions, children were asked why they thought the puppets took the gemstones.

Judgments. Then, to see whether children understood the transgression as morally wrong, and to examine whether they would judge this differently between conditions, M asked them to judge the severity of transgression, asking, "Do you think what they did was good, okay, not so good, or bad?" and "Why do you think this was [children's response]?"

Punishment. After that, to see whether children thought there should be consequences for the transgression, and whether their perception of these would differ between conditions, children were asked, "What should we do with them when they come back?" and "Should they be punished?" If children responded with "yes" to the latter question, they were then asked, "How should they be punished?"

Accountability. To see whether children were aware that tattling might have negative consequences for themselves if the transgressors found out that children had tattled, all children who had tattled were asked, "Would it be ok if they [the transgressors] knew that you told me about this?"

Exit choice and exit choice justification. After that, children were given a chance to leave their group: M discovered one additional group marker from the child's outgroup, and asked children whether they wanted to change groups or wanted to stay in their group, and why they decided to stay or leave.

Preference test. All four puppets, the child, and M sat together around the marble run for a final preference test. The child was given the first marble to put down the marble run. Then, M asked the child to distribute four more marbles, one by one, to the puppets.

After that, all the puppets and the child engaged in free play with the marble run.

Coding and Reliability

Along with exploring children's responses to the post-test questions, we also took a closer look at children's tattling responses. For the post-test questions, if children did not respond, shrugged, replied with "I don't know," or gave a response that was unrelated to the question, their answer was coded as *no answer*. Children were not prompted again in these cases, as these questions only aimed at providing

additional information. An independent, naïve coder coded a random sample of 25% of children's transcribed responses.

Tattling words. First, we looked at what words children used when tattling about the puppets' transgression. We were interested in whether children would use neutral vs. more negative words when tattling, as this could reflect their evaluation of the severity of the transgression. Thus, for all the children who tattled, we categorized the terms they used into *neutral* words, which described the event without evaluating it (e.g., "leaving only one," "taking out," "putting into her pocket") and more *negative* words (e.g., "stealing," "taking away," "swiping"). Reliability (Cohen's weighted kappa) was good with $\kappa = 0.69$.

Transgression justifications. After the main test, the moderator asked children why they thought the puppets had taken the gemstones. To investigate whether children would downplay vs. emphasize the transgression, their responses were coded as *positive/mitigating* (e.g., "because the stones were so pretty," "the stones were just lying around," "they were just curious"), *negative/aggravating* (e.g., "because they wanted to steal them," "because they are thieves"), and *neutral/other* (e.g., "they wanted to," "because they are in the yellow group"). Reliability was excellent with $\kappa = 1$.

Judgments and judgment justifications. Children's judgments of the puppets' transgression were coded as scores on a Likert scale ranging from 1 ("good") to 4 ("bad"). Children were also asked to justify their judgments. These justifications were coded into categories referring to *moral* (focusing on the harm that was done, e.g., "now only one is left," "if it belongs to someone, he will be sad when he returns"), *normative* (referring to norms, rules, permission, and ownership rights, e.g., "one shouldn't do this," "they didn't ask"), and *other* reasons (e.g., "because they were red," "I didn't think they would do it"). Reliability was excellent for children's judgments (Cohen's weighted kappa) with $\kappa = .94$, and fair for children's judgment justifications (Cohen's kappa) with $\kappa = 56$.

Punishment. We coded whether children thought that the transgressors should be punished ("yes") or not ("no"), and, if so, what kind of punishment they proposed in response to the other two questions (i.e., "What should we do with them?" and "How should they be punished?"). Since there was considerable overlap between children's responses to these similar questions, we combined their responses to these questions and, if they differed, counted the most severe form of punishment they suggested. Response categories were *lecturing* (verbal forms of punishment and lecturing, e.g., "scold them," "tell them that this is not okay"), *compensate* (requests to return the spoils, e.g., "they have to give the stones back"), or *penalty* (responses involving every other kind of non-verbal punishment, e.g., "send them to prison," "ban on watching TV," "catch them with a trap"). Reliability (Cohen's kappa) for all these measures together was excellent with $\kappa = .91$.

Accountability. We coded whether children thought that it would be okay if the transgressors knew about their tattling ("yes") or not okay ("no"). Reliability (Cohen's kappa) was excellent with $\kappa = 1$.

Exit choice and justification. For the exit choice, we assessed whether children wanted to leave their group or not ("yes" or "no"), and what reasons they gave for their choice. For their reasons, we coded the following categories: *transgression* (when children referred to the transgression, e.g., "because they took the gemstones," "because they didn't steal"), *policing* (when children wanted to stay in their group as a guard, e.g., "because I can tell you if they do something wrong," "I want to make sure that they will not steal again"), *group preference* (when children indicated that they identified with the group or preferred the group members in question, e.g., "I like the green group," "because I want to play with them"), *color preference* (when children referred to the group's color, e.g., "because green is my favorite color," "I like the yellow scarf better"), or *other* (any other statement that did not fall into one of the categories above, e.g., "because I want to"). Reliability (Cohen's kappa) was excellent with $\kappa = .91$.

Preference test. We assessed which puppet children gave the first marble to in the marble run game. If children gave this marble to one of their ingroup members, this was counted as preferring the ingroup, and if children gave this marble to one of the outgroup members, this was counted as preferring the outgroup. For children who had indicated that they wanted to change groups, we coded their preference according to their original group membership, i.e., the one they were allocated to at the beginning of the procedure. Reliability (Cohen's weighted kappa) was excellent with $\kappa = .92$.

Results

We did not push children to answer if they did not respond to the exploratory post-test questions. As a consequence the number of *no answer* responses was relatively high and the results should be taken with some caution. Thus, for most of the measures, statistical analysis was not appropriate due to small and uneven sample sizes in the different cells. In these cases we only report descriptive results.

Tattling words. We investigated whether children who tattled at some point ($n = 79$) used more negative words when tattling about the outgroup's transgression and/or the severe transgression. Table 1 shows the percentage of children who used each type of utterance in each condition. A Poisson-distributed GLM was run with group membership and transgression type as predictors. The full model did not differ from the null model ($p = .213$), indicating that there were no significant differences between conditions.

Transgression justifications. In response to the question about why the puppets had taken away the gemstones, children's relevant responses were mostly *positive/mitigating* (26%) or *neutral* responses (19%) overall. Table 1 shows the percentage of children who gave each response type in each condition. Any differences between conditions are difficult to interpret because of the small number of children who gave answers: overall, 50% of children did not answer this question.

Judgments and judgment justifications. Most children judged the transgression as either "not so good" or "bad." Only two children (both in the ingroup mild condition) judged the transgression as "good" or "ok". A Poisson-distributed GLM was run with group membership and transgression type as predictors for

children's judgment score, but the full model did not differ from the null model ($p >$ .25). Of the children who judged the transgression as being negative (i.e., either "not so good" or "bad"; all but two of the children), overall, 49% justified their responses with *normative* and 30% with *moral* reasons (see Table 1 for the justifications in each condition).

Punishment. Overall, 56% of children thought that the transgressors should be punished. Fourteen percent of all children suggested lecturing or scolding the transgressors verbally, 19% suggested that the transgressors be made to return the stones, and 25% proposed some sort of penalty. Strikingly, penalty, the most severe form of punishment, was rarely seen in the ingroup mild condition. Statistically there was no difference between conditions; the full-null GLM comparison was not significant ($p > .25$). However, again, because of high rates of *no answer* responses, this analysis should be interpreted with caution. Percentages for each condition are displayed in Table 1.

Accountability. Of the 79 children who tattled at some point, overall, 58% did not want the transgressors to know about their tattling, while 22% said it would be okay (20% gave no answer). Very few children in the ingroup mild condition wanted the transgressors to know about their tattling (see Table 1). However, statistically this difference between conditions was only marginal: The null model differed only marginally from the full model ($\chi^2(3) = 6.57, p = .087$). The full model revealed a marginal interaction effect (*Estimate* = -2.46, *SE* = 1.39, $\chi^2(1) = 3.67$, $p = .055$, Nagelkerke's $R^2 = .08$). Pairwise comparisons revealed marginal differences between the ingroup mild condition and all other conditions (Fisher's exact tests: $p = .067$, risk ratio = 6.07 with the ingroup severe condition; $p = .085$, risk ratio = 6 with the outgroup mild condition, and $p = .085$, risk ratio = 0.88 with outgroup severe condition).

Exit choice and exit choice justifications. When children's ingroup members were the transgressors, 46% of children wanted to leave their group after the mild transgression, but only 25% after the severe transgression. When the transgressors were outgroup members, 17% wanted to leave their group (and join the transgressor group) after the mild transgression, and 8% after the severe transgression (Table 1). Because children's tattling behavior might have influenced children's wish to change groups, we first tested whether children's exit choice was influenced by their tattling (yes/no), but found no relation between tattling and children's exit choice, GLM full vs. null model, $p >.25$). Then, a GLM was run to analyze the effect of condition on children's exit choice. The full model differed significantly from the null model ($\chi^2(3) = 10.22$, $p = .017$) and revealed no interaction between group membership and transgression type ($p > .25$). A main effect of group membership was found (*Estimate* = 1.39, *SE* = .54, $z = 2.57, p = .0091$, Nagelkerke's $R^2 = .11$), as well as a marginal main effect of transgression type ($p = .08$). Thus, children were more likely to leave their group when their group members had transgressed, and marginally less likely to leave their group after a severe transgression.

Overall, most children justified their response with their *color preference* (26%), a *group preference* (21%), the *transgression* (18%) or *other* reasons (19%). Only four percent of children justified their response with *policing*. Table 2 shows the percentage of children's reasons by condition. The transgression type did not

seem to have a big impact on children's justification. We additionally looked at the data clustered in the following way: First, we were interested in what reasons children gave for *not wanting to be a member of the transgressors' group*, that is, why children in the ingroup condition wanted to leave the transgressors' group ($n = 17$), and why children in the outgroup condition did not want to join the transgressors' group ($n = 42$). Taken together, 29% of these children justified their decision with the *transgression* they had observed. Other common reasons were *color preference* (22%) and *group preference* (19%). Conversely, we were also interested in what reasons children gave for *wanting to be a member of the transgressors' group*, and therefore looked at the justifications of children in the ingroup condition who wanted to stay in the transgressors' group ($n = 31$) and children in the outgroup condition who wanted to join the transgressors' group ($n = 6$). Taken together, these children mostly explained their choice with *color preference* (32%) or *group preference* (24%). Interestingly, three of the children (8%) who wanted to stay in the transgressors' group justified this with *policing*, indicating that they would like to improve their group's behavior from within.

Preference test. Finally, we assessed children's ingroup preference by looking at whom they gave the first marble to when they were asked to allocate the marbles for the marble run game. In the ingroup condition, only 46% of children preferred their ingroup member in this task, compared to 69% of children in the outgroup condition who preferred their ingroup (see Table 1 for percentages by group and transgression conditions). To investigate the effect of condition on children's ingroup preference, we ran a GLM with group membership and transgression type as predictors for this binomial measure of ingroup preference. The null model differed only marginally from the full model ($\chi^2(3) = 6.65$, $p = .084$), and the interaction was not significant ($p > .25$). After removing the interaction from the model ($\chi^2(2) = 6.33$, $p = .042$), we found a main effect of group membership (*Estimate* $= 0.97$, *SE* $= 0.43$, $z = 2.26$, $p = .024$, Nagelkerke's $R^2 = .07$). Thus, children's ingroup preference was diminished when the transgression was conducted by their ingroup, compared to outgroup, members, while the severity of the transgression did not influence children's ingroup preference.

Tables

Table 1

For the Post-Test Measures, the Percentage of Children Giving each Type of Answer in each Group Membership and Transgression Type Condition

Measures	Ingroup Conditions		Outgroup Conditions	
	Mild	Severe	Mild	Severe
Tattling Words (*n* = 79)				
Negative	48%	69%	74%	61%
Neutral	52%	31%	26%	39%
Transgression Justifications (*n* = 96)				
Positive/Mitigating	33%	29%	17%	25%
Neutral	29%	13%	17%	17%
Negative/Aggravating	0%	4%	13%	4%
No answer	38%	54%	54%	54%
Judgments (*n* = 96)				
Good	4%	0%	0%	0%
Okay	4%	0%	0%	0%
Not so good	46%	46%	38%	29%
Bad	46%	42%	58%	67%
No answer	0%	13%	4%	5%
Judgment Justification (*n* = 94)				
Moral	23%	21%	33%	42%
Norm	50%	54%	58%	33%
Other	5%	0%	0%	4%
No answer	23%	25%	8%	21%
Punishment (*n* = 96)				
Yes	58%	54%	58%	54%
No	29%	29%	21%	21%
No answer	13%	17%	21%	25%
How to punish (*n* = 96)				
Lecturing	13%	17%	17%	8%
Compensating	29%	8%	13%	25%
Penalty	8%	25%	33%	33%
No answer	50%	50%	38%	33%
Accountability (*n* = 79)				
Yes	5%	31%	32%	22%
No	76%	56%	53%	48%
No answer	19%	13%	16%	30%
Exit Choice (*n* = 96)				
Leaving	46%	25%	17%	8%
Preference Test (*n* = 96)				
Ingroup Preference	38%	54%	67%	71%

Note. Justification judgments include only children who judged the transgression as negative (i.e., all but two children).

Table 2
Reasons Children in Each Group Membership and Transgression Type Condition Gave for Changing Groups or Staying in Their Group

Reasons	Ingroup Conditions		Outgroup Conditions	
	Mild	Severe	Mild	Severe
Changing (*n* = 23)				
Transgression	27%	33%	0%	0%
Policing	0%	0%	0%	0%
Group Preference	27%	17%	25%	50%
Color Preference	27%	17%	50%	50%
Other	9%	33%	0%	0%
No Answer	9%	0%	25%	0%
Staying (*n* = 73)				
Transgression	0%	0%	25%	32%
Policing	15%	6%	5%	0%
Group Preference	31%	17%	15%	18%
Color Preference	23%	33%	20%	23%
Other	23%	28%	20%	14%
No Answer	8%	17%	15%	14%