

THIS MANUSCRIPT HAS NOT BEEN PEER-REVIEWED

The Cognitive Processes Underlying Moral Judgment Across Development

Lisa Chalik¹, Jay Van Bavel², & Marjorie Rhodes²

¹Yale University

²New York University

Address all correspondence to: Lisa Chalik; lisa.chalik@yale.edu; Yale University Department of Psychology, 2 Hillhouse Avenue, New Haven, CT 06511

Abstract

Some moral philosophers have suggested that a basic prohibition against intentional harm ought to be at the core of moral belief systems across human societies. Yet, experimental work suggests that not all harm is created equal—people often respond more negatively to harm that occurs among fellow social group members, rather than between members of different groups. The present two studies investigated how concerns about social group membership factor into the moral judgment system. Adults ($N = 111$, Study 1) and children ($N = 110$, Study 2) evaluated instances of inter- and intra-group harm under varying levels of cognitive load. Both children and adults responded more slowly to intergroup harm than to intragroup harm. Furthermore, adults under cognitive load rated intergroup harm more leniently than intragroup harm, but adults who were not under load rated the two types of behaviors similarly. These findings suggest that across development, evaluations of intergroup harm rely more heavily on conscious deliberation than evaluations of intragroup harm. Thus, people's evaluations of harmful behaviors are made in light of information about the social category membership of the people involved.

Keywords: moral judgment; social cognitive development; intergroup cognition

Introduction

In the Fall of 2009, an army major on a United States military base in Fort Hood, Texas opened fire on his fellow American soldiers, killing 13 people and injuring more than 30 others. As with other mass shootings, this tragedy evoked outrage around the country. Yet, this particular event was quite different from other mass shootings. It was not a random attack upon people mostly unknown to the attacker—it was an attack by an American soldier upon fellow members of the American military. In a statement at a memorial service for the victims, President Obama stated that the fact that the attack was committed by an American, on American soil, “makes the tragedy even more painful, even more incomprehensible” (The White House, 2009).

This response illustrates a common phenomenon: People often treat harm against members of the *same* social group (e.g., among fellow Americans) as more morally reprehensible than harm between members of *different* groups. Yet, the question of how concerns for social groups fit into the moral judgment system is a challenging one. Many moral philosophers and psychologists have argued that the prohibition against intentional harm is the most fundamental component of human morality, regardless of other features of a situation, such as group membership (Nagel, 1987; Smetana, 2006). By this account, group membership should not factor into people’s initial evaluation of a harmful interaction—people should simply respond negatively to any sort of harm. Indeed, a general prohibition against harming others exists across human cultures (Haidt & Joseph, 2004; Helwig, 2006; Wainryb, 2006).

Yet, experimental work has revealed that not all harm is viewed equally. People often evaluate harm more or less severely depending on features of the situation aside from the harmful behavior itself, such as group membership (Meier & Hinsz, 2004; Rai & Fiske, 2011;

Rozin, Lowery, Imada, & Haidt, 1999; Shweder, Much, Mahapatra, & Park, 1997). For example, across many human societies, people actually value violence toward out-group members, especially if they identify strongly with their in-group (Cohen, Montoya, & Insko, 2006), and viewing harm against out-group members is associated with the activation of brain regions that encode reward (Cikara, Botvinick, & Fiske, 2011). Thus, it remains an open question exactly how the moral judgment system weighs concerns for group membership against more general prohibitions against harm. Do people arrive at their moral evaluations by first focusing only on harm, then updating their evaluations after considering other aspects of a situation (e.g., group membership)? Or are people's *initial* evaluations of harmful events made in light of information about the group membership of those involved?

Developmental research provides one possible answer to the question of how group membership factors into the moral judgment system. Young children have a basic expectation that social groups mark people who hold intrinsic, moral obligations toward one another. For instance, one study introduced 3- to 9-year-old children to two novel social categories and showed them instances of intra- or inter-group harm (e.g., someone teasing a member of their own or another group; Rhodes & Chalik, 2013). Children were asked to give an initial evaluation of the harmful action. Then, children were told that there were no explicit rules prohibiting the harmful behavior in the actors' context (e.g., there were no rules prohibiting teasing), and were asked to evaluate the action a second time. When the actions involved members of the same group, children rated the action as equivalently bad regardless of the rule, suggesting that they thought the agent was intrinsically obligated not to harm a member of the agent's own group, regardless of the circumstance. In contrast, when the actions involved members of different

groups, children evaluated the action more leniently after they were told there were no specific rules in place prohibiting the behavior. This pattern of findings suggests that children have a basic intuition that people should not harm their own group members, but require more information (e.g., consideration of local rules) to determine when intergroup harm is prohibited or permissible. No prior research, however, has investigated the role of cognitive resources in these judgments.

Based on these findings, as well as prior work to suggest that moral judgments vary in the degree to which they rely on deliberative reasoning (Greene et al., 2001), the present work tests the hypothesis that deliberative reasoning plays a more critical role in the evaluation of intergroup harm, whereas the evaluation of intragroup harm is more reflexive. In other words, a prohibition against intragroup harm (but not intergroup harm) is present in the expectations that children hold very early in life. Because intergroup harmful behaviors do not violate this prohibition, evaluating these behaviors as wrong requires more conscious deliberation (e.g., consideration of local rules and other contextual features). On the other hand, behaviors that do violate the basic prohibition against intragroup harm (i.e., harmful actions among members of the same group) should be reflexively judged as wrong, and require less deliberation. To test these hypotheses, we recorded adults' and children's response times to and ratings of negative intra- and inter-group behaviors under varying levels of cognitive load.

We made two predictions: First, if reactions against intergroup harm require more deliberation than those against intragroup harm, then people should be slower to evaluate intergroup harm than they are to evaluate intragroup harm. Second, since deliberative responses require more cognitive effort than responses that rely more heavily on intuition (Chaiken &

Trope, 1999; Kahneman, 2003; Stanovich & West, 2000), cognitive load should interfere more severely with reasoning processes that require more deliberation. Thus, our second prediction was that people would evaluate intergroup harm, but not intragroup harm, more leniently under cognitive load.

Study 1

Participants

We recruited 111 adults (M age = 20.4, range = 19.0 - 24.0, 82 female) from New York University in exchange for course credit. Data were collected during the 2015-2016 academic year. An additional 5 participants were tested but excluded from analysis because they failed to correctly answer attention check questions. Participants were randomly assigned to the No Load ($n = 55$) or Load ($n = 56$) conditions.

Procedure

Participants sat in front of a computer. On the screen, they saw a seven-point scale ranging from -3 to 3, represented visually by smiley faces (-3 = big frown, 0 = neutral face, 3 = big smile). Participants were told that they would see a series of social interactions, and for each, they had to choose the corresponding point on the scale. Participants completed two blocks of four practice trials each to familiarize them with this procedure. For each trial, a social interaction appeared (as a hand-drawn picture) and was described in an audio recording, and the participant had to press a key corresponding to one of the points on the scale (denoted by images on the computer keys).

For participants in the No Load condition, the second block of practice trials was a repeat of the first block. For participants in the Load condition, the second block of practice trials included a cognitive load manipulation: In addition to rating each interaction, participants performed a trial-by-trial digit span memory task (Longstaffe, Hood, & Gilchrist, 2014). In each trial, a 5-digit number appeared on the screen for three seconds, followed by the social interaction. After rating the interaction, participants were prompted to enter the string of digits that they had seen before moving to the next trial. We used this cognitive load manipulation because we hypothesized that when people deliberate over moral evaluations, their deliberation centers around retrieving the relevant social rule (or rules) from their semantic memory and deciding whether it applies in the present scenario; for example, when thinking about stealing, an individual might first think about whether there exist social rules against stealing, then think about whether the present behavior violates those rules. Thus, any manipulation intended to interfere with this process would need to interfere directly with the retrieval of semantic knowledge. Because the working memory system acts to control attention and allow for the retrieval of information (Engle, 2002), we predicted that a digit span memory task, which taxes the working memory system, would interfere with participants' evaluations of intergroup harm.

After the practice trials, participants read a short story on the screen in which they were introduced to two novel groups of children, marked by shirt color and team names (the "Flurps," wearing blue shirts, and the "Zazzes," wearing red shirts). The groups were described as engaged in noncompetitive activities—each group was working together to build a tower out of blocks. Participants completed two attention check questions ("Look at these two children. Are they in the same group or different groups?") to ensure that they recognized the two groups. Next,

participants completed 12 test trials in which they saw and heard about a social interaction that had occurred among the characters in the story, then evaluated the interaction by choosing a point on the scale. Six trials were about interactions that had occurred among members of the same group (e.g., between a Flurp and another Flurp), and six trials were about interactions that had occurred between members of different groups (e.g., between a Flurp and a Zazz). The intragroup and intergroup interactions were presented in blocked counterbalanced order. Additionally, for each group context, half of the social interactions were about harmful behaviors (intended to test our hypotheses about the cognitive processes underlying judgments about inter- and intra-group harm) and half of the social interactions were about prosocial behaviors (intended as control items to ensure that participants used the full range of the scale). In the Load condition, the test trials included the digit span memory task described above.



Figure 1. Example of one trial in the Load condition. Participants saw a 5-digit number, then saw and heard about a social interaction. After rating the interaction (by pressing a key corresponding to one of the smiley faces), participants were prompted to enter the number they had seen. In the No Load condition, participants gave the same ratings without the digit-span memory task.

We recorded participants' ratings of the behaviors as well as their reaction times for each rating. Reaction times were measured from the moment the interaction appeared on the screen until the moment participants chose a point on the rating scale. Ratings are presented as participants' average rating of how bad the harmful behaviors were, with higher numbers

indicating a more negative rating. Reaction times were log-transformed for analysis, but for ease of interpretation are presented in milliseconds. Trials in which participants took fewer than 500 ms or more than 10,000 ms to respond were excluded from analysis. All data and code can be found on the Open Science Framework at <https://osf.io/xr2wh/>.

Results

To test our first hypothesis, that people would take longer to evaluate intergroup harm than they would to evaluate intragroup harm, we averaged participants' reaction times for the three harmful social interactions in each group context (intergroup and intragroup). We subjected these times to a repeated measures ANOVA with cognitive load (Load or No Load) as a between-subjects factor and group context (intergroup or intragroup) as a within-subjects factor. We found a main effect of group context, such that participants responded more slowly for intergroup harm ($M = 3899$, $CI = 3700 - 4098$) than they did for intragroup harm ($M = 3297$, $CI = 3150 - 3444$), $F(1,109) = 30.40$, $p < .001$, partial $\eta^2 = .22$. There was no main or interactive effect of cognitive load.

To test our second hypothesis, that cognitive load would selectively interfere with participants' ratings of intergroup, but not intragroup, harm, we repeated the above analysis on participants' ratings, rather than their reaction times. We found a main effect of group context, indicating that participants rated intragroup harm as worse than intergroup harm, $F(1,109) = 4.52$, $p < .05$, partial $\eta^2 = .04$. We also found a main effect of cognitive load, indicating that participants in the Load condition rated behaviors as worse than participants in the No load condition, $F(1,109) = 5.32$, $p < .05$, partial $\eta^2 = .05$.

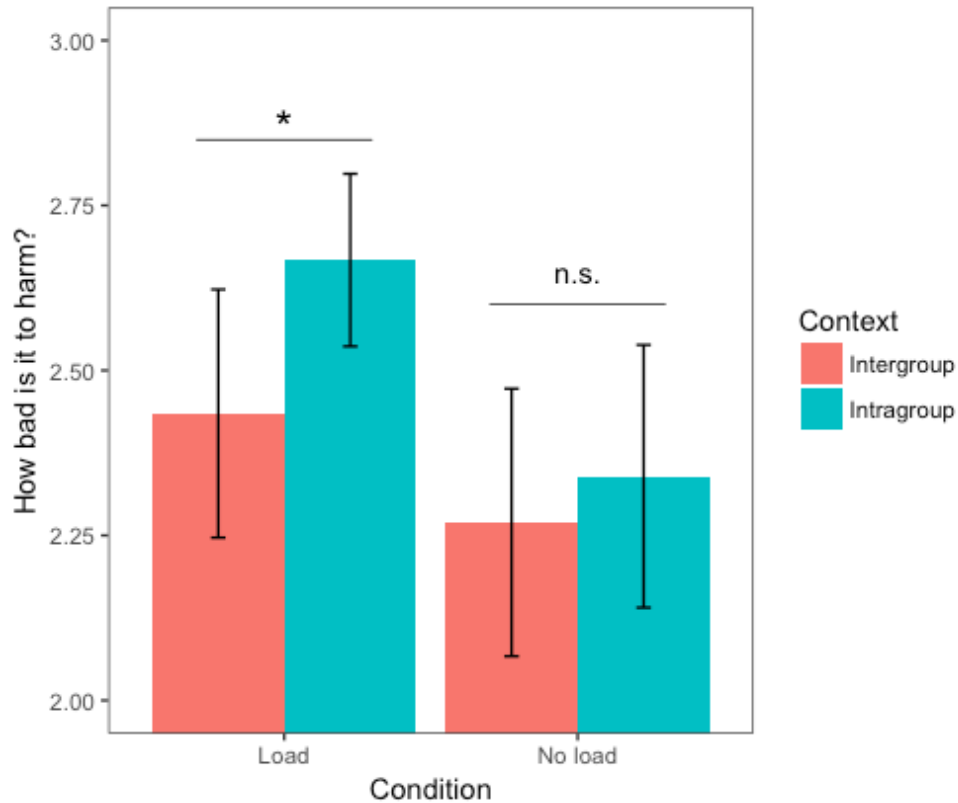


Figure 2. Ratings of inter- and intra-group harm in each condition. Error bars represent 95% confidence intervals.

Contrary to our prediction, the interaction between group context and cognitive load was not statistically significant ($F(1,109) = 1.29, p = .26$), but an examination of the means suggests that the effect of context was driven by participants in the Load condition (see Figure 2). In the Load condition, participants reliably evaluated intragroup harm ($M = 2.67, CI = 2.50 - 2.83$) as worse than intergroup harm ($M = 2.44, CI = 2.24 - 2.63$), $t(55) = 2.51, p < .05$, Cohen's $d = .34$, whereas in the No Load condition, participants evaluated the two types of harm similarly (intragroup: $M = 2.34, CI = 2.17 - 2.51$; intergroup: $M = 2.27, CI = 2.08 - 2.46$), $t(54) = .64, p = .53$, Cohen's $d = .09$.

To investigate these effects further, as a set of post-hoc analyses, we tested whether the above effects differed based on the extent to which participants' cognitive resources were taxed in the Load condition: Participants in this condition correctly reported the number that they had been told to remember on an average of 4.6 trials (out of the six trials that involved harmful behaviors). Because this task was designed to interfere with the relevant cognitive processing for these types of moral evaluations, participants who were more successful on the task (i.e., who remembered more numbers correctly) should have had less working memory capacity available to allow them to deliberate over their evaluations. We thus tested the correlation between performance on the memory task and the difference between participants' ratings of inter- and

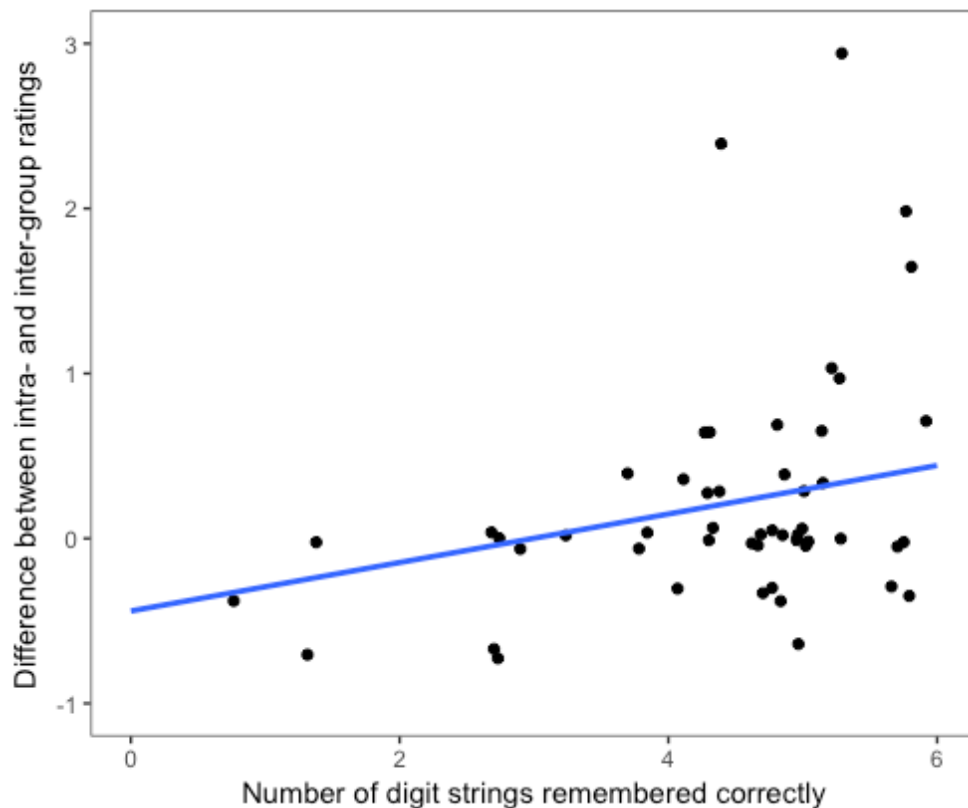


Figure 3. The difference between participants' ratings of intra- and inter-group harm, by the number of digit strings that they correctly remembered. Participants who remembered more digit strings (i.e., who were more taxed by the cognitive load manipulation) showed a greater difference, rating intragroup harm as worse than intergroup harm.

intra-group harm. The difference between participants' ratings was positively correlated with memory performance ($r = .26, p < .05$; see Figure 3), such that participants who remembered more numbers correctly showed a greater difference between their ratings of intra- and inter-group harm (rating intragroup harm as worse than intergroup harm). This finding suggests that a stronger load manipulation might produce an even stronger pattern of results.

Discussion

Across both of the measures that we tested (ratings and reaction times), participants responded differently for intra- and inter-group harm. In both conditions, participants responded more slowly for intergroup harm than for intragroup harm; this difference was quite large, suggesting that evaluations of intergroup harm require more conscious deliberation than evaluations of intragroup harm. Also, participants rated intragroup harm as worse than intergroup harm. Furthermore, cognitive load interfered modestly with evaluations of intergroup harm—participants under load rated intergroup harm as less bad than intragroup harm, whereas participants under no load evaluated the two types of harm similarly (although the interaction effect was not statistically significant). Additionally, a post-hoc analysis revealed that the difference between ratings of intergroup and intragroup harm among participants under cognitive load was more pronounced for participants whose cognitive resources were more taxed by the load manipulation. Taken together, these findings suggest that evaluations of intragroup harm rely more heavily on reflexive responses, whereas evaluations of intergroup harm rely more heavily on conscious deliberation.

Study 2

In Study 2, we tested whether the effects we documented in adults are continuous throughout the lifespan by conducting a similar study with 4- to 6-year-old children. No prior research has investigated the role of cognitive resources in children's evaluations of inter- and intra-group harm. It is possible that these processes are different from those we found in adults; although young children certainly use social categories when evaluating harmful behaviors (Rhodes & Chalik, 2013), it is possible that the adult evaluations—evoking more or less deliberation in different group contexts—emerge slowly over time. On this account, we would expect deliberative responses, which have been documented in young children in prior research on moral development (Smetana, 1985), to play a consistent role in children's evaluations across various types of behaviors.

It is also possible that the cognitive processes that guide evaluations of inter- and intra-group harm in adults are similarly variable in young children. Prior work supports this possibility; as early as the preschool years, children treat intragroup harm as a serious moral violation, whereas they treat intergroup harm as wrong for more conventional reasons (Rhodes & Chalik, 2013). Thus, there does appear to be some variation in how children arrive at their judgments of inter- and intragroup harm. Yet, no work has examined the cognitive processes underlying these judgments. If children's evaluations are driven by the same variation in deliberation that we found in adults in Study 1, then we should find similar effects in children's reaction times to and their ratings of inter- and intra-group harm.

Participants

We recruited 110 4- to 6-year-old children (M age = 5.35, range = 3.98 - 7.07, 57 female) at the Children's Museum of Manhattan. Researchers approached parents at the museum and invited them to participate in research. Once parents had given consent, children participated in a quiet room at the museum. An additional 9 children were tested but excluded from analysis because they did not complete the entire testing session. Children were randomly assigned to the No Load ($n = 52$) or Load ($n = 58$) conditions.

Procedure

Children performed the same task as the adults in Study 1 with the exception that instead of a digit span recall task, in the Load condition, we used a prospective memory task—requiring participants to remember to perform a future action after a cue. Prospective memory tasks have been shown to interfere with performance in ongoing tasks in 4- to 6-year-old children (Leigh & Marcovitch, 2014). Thus, in the Load condition, children were told that they had to look for a picture that contained someone wearing green shoes, and they were instructed to ring a bell when they saw this special picture. Children saw two prospective memory trials (pictures containing green shoes) throughout the study—one after each block of six test trials. Thus, children saw 14 trials total (six intragroup test trials, one intragroup prospective memory trial, six intergroup test trials, one intergroup prospective memory trial), counterbalanced in the same manner as in Study 1. Because the prospective memory trials required a different type of response from the test trials, these two trials were not included in our analyses—thus, our analyses included only the six

test trials in each block. Children therefore completed the same number of test trials as adults in Study 1 (who did not see prospective memory trials).

As in Study 1, we recorded participants' ratings of the behaviors as well as their reaction times for each rating. Trials in which participants took fewer than 500 ms or more than 10,000 ms to respond were excluded from analysis. An additional two children who took longer than 10,000 ms to respond on over 25% of trials were excluded from analysis.

Results

To test whether children took longer to evaluate intergroup harm than intragroup harm, we subjected children's average reaction times to a repeated measures ANOVA with condition (Load or No Load) as a between-subjects factor and group context (intergroup or intragroup) as a within-subjects factor. We replicated the main effect of group context, indicating that children responded more slowly for intergroup harm ($M = 5344$, $CI = 5108 - 5579$) than they did for intragroup harm ($M = 4894$, $CI = 4600 - 5188$), $F(1,108) = 11.00$, $p < .005$, partial $\eta^2 = .09$. There was no main or interactive effect of condition.

To test whether cognitive load interfered with children's ratings of the behaviors, we repeated the above analysis with rating as the dependent variable. There were no significant effects, indicating that children responded similarly for intragroup and intergroup harm in both conditions (Load: intragroup $M = 2.10$, $CI = 1.74 - 2.45$, intergroup $M = 2.26$, $CI = 1.99 - 2.54$; No load: intragroup $M = 2.19$, $CI = 1.84 - 2.53$, intergroup $M = 2.02$, $CI = 1.71 - 2.32$; $ps > .13$). We then investigated whether the effects differed based on the extent to which children engaged in the prospective memory task: Of the 58 children in the Load condition, 25 successfully remembered to ring the bell on the first prospective memory trial, and 27 failed to do so (these

children needed prompting to remember that they were supposed to ring the bell). The remaining six children's sessions were not videotaped, and we thus could not code whether they performed the prospective memory task. We then repeated the above analysis excluding children who failed the prospective memory task. Again, there were no significant effects ($ps > .41$), indicating that children responded similarly for intragroup and intergroup harm in both conditions. Thus, the lack of an effect of cognitive load here may be due to the fact that the manipulation was not strong enough to divert children's cognitive resources.

Discussion

We replicated our reaction time results from Study 1 in young children—children responded more slowly for intergroup harm than for intragroup harm. But contrary to our findings with adults, we found no effect on children's ratings of the behaviors—children rated the harmful behaviors as equivalently bad, regardless of condition.

There are several possible explanations for these discrepant findings across the two measures (ratings and reaction times). One possibility is that, as we hypothesized, children were slower to evaluate intergroup harm than intragroup harm because they use more deliberative processing to evaluate intergroup harm. From this perspective, the cognitive load manipulation (the prospective memory task) should have selectively interfered with children's evaluations of intergroup harm, as it did among adults in Study 1. Yet, perhaps the load manipulation that we used here was not successful among young children. Similar manipulations have been used with children of these ages in previous work (Leigh & Marcovitch, 2014). However, the present task was considerably more complicated than previous studies. Thus, perhaps children attended only to the evaluation task, because it was more complicated, and did not devote any cognitive

resources to the prospective memory task. If so, then children in the Load condition did not experience increased cognitive load, despite the additional task that we asked them to perform. The fact that excluding children who failed the prospective memory task did not alter our findings lends support to this possibility. If this is the case, then a stronger load manipulation may selectively interfere with children's evaluations of intergroup harm.

Alternately, perhaps children were slower to evaluate intergroup harm not because evaluating these behaviors requires increased deliberation, but because of some other feature of the intergroup trials. For example, these trials may have required children to process more information than the intragroup trials did (i.e., children had to note the presence of two category memberships, instead of just one), so perhaps this additional processing demand in the intergroup trials caused children to respond more slowly, but not for any reason that had to do with their evaluations of the harmful behaviors. If this is the case, we might expect that children's evaluations of intergroup harm would be unaffected by an even stronger cognitive load manipulation. Future research should distinguish these possibilities.

General Discussion

The present studies represent the first evidence that, among adults, evaluations of intergroup harm rely more on deliberative responses than evaluations of intragroup harm. Two pieces of evidence support this claim. First, adults were slower to evaluate intergroup harm than intragroup harm. Second, under load, adults evaluated intergroup harm more leniently than they evaluated intragroup harm, whereas under no load, they evaluated the two types of harm similarly. This pattern suggests that evaluations of intergroup harm may require more deliberation. There was some evidence for developmental continuity in these effects—young

children were also slower to evaluate intergroup than intragroup harm. Yet, because children's evaluations were unaffected by the cognitive load manipulation, it remains unclear whether similar processes shaped children's and adults responses to these scenarios.

These studies provide evidence that different cognitive processes underlie the moral judgments that occur in different types of intergroup contexts. A great deal of prior research has shown that concerns for social groups are an important part of moral codes across human cultures (Haidt & Joseph, 2004; Meier & Hinsz, 2004; Rai & Fiske, 2011; Rozin et al., 1999; Shweder et al., 1997), yet no prior work has attempted to document the actual processes by which beliefs about social groups operate when people evaluate moral scenarios. In the present studies, we have shown that a combination of reflexive and deliberative processes shapes people's responses to these types of scenarios. Specifically, we have shown that scenarios that do not violate people's basic intuitions about the function of social categories (i.e., harm between people from different groups) evoke more conscious deliberation than scenarios that do violate those intuitions (i.e., harm among people from the same group).

These findings have important implications for the study of moral development. Some theories have posited that moral evaluation exists as a separate domain from other types of reasoning. On this account, people make moral judgments on the basis of whether an action causes harm or unfair treatment, and only after the initial judgment is made do they incorporate considerations for parts of the scenario aside from the behavior itself, such as group membership (Killen & Rizzo, 2014; Killen, Rutland, Abrams, Mulvey, & Hitti, 2013). On this account, when people see a harmful behavior occur, they immediately evaluate the behavior negatively because it violates their basic intuition that harm should be avoided. Then, after they have generated an

initial evaluation, they update that evaluation on the basis of information about social group membership and other relevant factors (possibly evaluating intragroup harm more harshly, and intergroup harm more leniently, than they initially had).

The present findings suggest an alternative account: that people's initial evaluations of harmful behaviors are made in light of information about the social category membership of the people involved. This is the intuition that President Obama invoked when he noted that an American shooting Americans on American soil made a tragedy even more painful.

Our results suggest that children hold the same basic expectations about the function of social categories very early in life. Their use of social categories in evaluating moral scenarios appears to be a very early-emerging feature of human cognition. Across both studies, we designed the stimuli and language to be simple enough that it could be used with both children and adults. This approach was helpful in that it allowed us to run similar studies with adults and with children, which strengthens the conclusions that we can draw about the relationship between children's and adults' moral evaluations. As such, the present work suggests that this part of the moral judgment system remains relatively stable across development. It is thus striking that we did find such a clear pattern of responses in adults while using child-friendly stimuli. Still, future work should investigate adults' responses to a wider range of complex intergroup scenarios.

Future work should also further investigate the processes underlying children's judgments of inter- and intragroup harm. As noted above, the present studies can be taken as some evidence for developmental continuity in these processes; both adults and children responded more slowly to intergroup harm than to intragroup harm, suggesting that across development, evaluations of

intergroup harm rely more on deliberative processing. Yet, because the cognitive load manipulation used here did not influence children's judgments, any conclusions that we are able to draw regarding developmental continuity remain tentative. To fully investigate whether children's evaluations are guided by the same underlying processes as those found in adults, future work should examine children's evaluations of inter- and intra-group harmful behaviors while putting them under a greater degree of cognitive load. If a stronger cognitive load manipulation selectively interferes with children's judgments in the same way that it interfered with those of adults, then we will be able to conclude more strongly that children, like adults, rely more heavily on deliberative processing when evaluating intergroup harm than when evaluating intragroup harm.

Despite these open questions, the present work represents an important contribution to the literatures on child and adult moral cognition. These studies have provided the first step toward documenting the cognitive processes that underlie people's responses to inter- and intra-group harm. In doing so, they have expanded our understanding of the nature of adult moral cognition as well as the processes that shape moral cognition across development.

Acknowledgments

We are very grateful to the families and staff at the Children's Museum of Manhattan, and to the members of the NYU Conceptual Development and Social Cognition Lab. Funding was provided by NSF BCS-1226942, NSF BCS-1147543, and NSF BCS-1349089.

References

- Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford Press.
- Cikara, M., Botvinick, M.M., & Fiske, S.T. (2011). Us versus them: Social identity shapes neural responses to intergroup competition and harm. *Psychological Science, 22*, 306-313.
- Cohen, T., Montoya, R., & Insko, C. (2006). Group morality and inter-group relations: Cross-cultural and experimental evidence. *Personality and Social Psychology Bulletin, 11*, 1559-1572.
- Engle, R.W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*, 19-23.
- Greene, J.D., Sommerville, R..B., Nystrom, L.E., Darley, J.M., & Cohen, J.D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*, 2105-2108.
- Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus: Special Issue on Human Nature, 133*, 55-66.
- Helwig, C.C. (2006). Rights, civil liberties, and democracy across cultures. In M. Killen & J.G. Smetana (Eds.), *Handbook of moral development*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58*, 697-720.
- Killen, M., & Rizzo, M. (2014). Morality, intentionality, and intergroup attitudes. *Behaviour, 151*, 337-359.
- Killen, M., Rutland, A., Abrams, D., Mulvey, K.L., & Hitti, A. (2013). Development of intra- and

- intergroup judgments in the context of moral and social-conventional norms. *Child Development, 84*, 1063-1080.
- Leigh, J., & Marcovitch, S. (2014). The cognitive cost of event-based prospective memory in children. *Journal of Experimental Child Psychology, 127*, 24-35.
- Longstaffe, K.A., Hood, B.M., & Gilchrist, I.D. (2014). The influence of cognitive load on spatial search performance. *Attention, Perception, & Psychophysics, 76*, 49-63.
- Meier, B., & Hinsz, V. (2004). A comparison of human aggression committed by groups and individuals: An interindividual-intergroup discontinuity. *Journal of Experimental Social Psychology, 40*, 551-559.
- Nagel, T. (1987). *What does it all mean?* New York: Oxford University Press.
- Rai, T.S., & Fiske, A.P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review, 118*, 57-75.
- Rhodes, M., & Chalik, L. (2013). Social categories as markers of intrinsic interpersonal obligations. *Psychological Science, 24*, 999-1006.
- Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology, 76*, 574-586.
- Shweder, T.A., Much, N.C., Mahapatra, M., & Park, L. (1997). The “big three” of morality (autonomy, community, divinity) and the “big three” explanations of suffering. In A.M. Brandt & P. Rozin (Eds.), *Morality and health*. Florence: Taylor & Frances/Routledge.
- Smetana, J.G. (1985). Preschool children’s conceptions of transgressions: Effects of varying

moral and conventional domain-related attributes. *Developmental Psychology*, 21, 18-29.

Smetana, J. (2006). Social domain theory: Consistencies and variations in children's moral and social judgments. In M. Killen & J.G. Smetana (Eds.), *Handbook of moral development*. Mahwah, NJ: Erlbaum.

Stanovich, K.E., & West, R.F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23, 645-665.

The White House, Office of the Press Secretary. (2009). Remarks by the President at memorial service at Fort Hood [Press release]. Retrieved from <https://www.whitehouse.gov/the-press-office/remarks-president-memorial-service-fort-hood>

Wainryb, C. (2006). Moral development in culture: Diversity, tolerance, and justice. In M. Killen & J.G. Smetana (Eds.), *Handbook of moral development*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.