

## Naïve Theories of Social Groups

Marjorie Rhodes

*New York University*

Four studies examined children's (ages 3–10, Total  $N = 235$ ) naïve theories of social groups, in particular, their expectations about how group memberships constrain social interactions. After introduction to novel groups of people, preschoolers (ages 3–5) reliably expected agents from one group to harm members of the other group (rather than members of their own) but expected agents to help members of both groups equally often. Preschoolers expected between-group harm across multiple ways of defining social groups. Older children (ages 6–10) reliably expected agents to harm members of the other group and to help members of their own. Implications for the development of social cognition are discussed.

To make sense of human behavior, children consider a wide range of unobservable causal factors. Imagine that a child sees a girl, Jane, carrying a doll across a room and smiling as she shows it to a friend. By the preschool years, children understand this sequence by appealing to abstract theories of human action (Wellman & Gelman, 1992). For example, children can rely on their naïve theories of psychology to infer that underlying mental states motivate Jane's behavior (e.g., that she *likes* the doll, that she *wants* to share her doll, and that she *believes* that her friend will like the doll too; Wellman, 2002). Children can also understand Jane's behavior by appealing to naïve theories of sociology—theories that reference social causes extending beyond the individual—and thus consider her category memberships and relevant social norms (e.g., that Jane and her friend will play with the doll because they are *girls*, and that perhaps there is a *social rule or expectation* that Jane will share her toys with other children; Gelman, Collman, & Maccoby, 1986; Kalish & Shiverick, 2004; Taylor, Rhodes, & Gelman, 2009).

Naïve theories are important because they allow children to make inferences that go beyond directly observable events and to infer how new people will behave in novel situations. Although young

preschoolers—and even infants (e.g., Woodward, 1998)—have basic theories of human action, these theories also develop and change across childhood. For example, whereas preschoolers often consider only goals or desires to predict behavior, older children attribute more weight to beliefs (Wellman, Cross, & Watson, 2001). Also, preschoolers view categories as directly determining behavior (e.g., being a girl *directly causes* a child to like dolls), whereas older children view this process as probabilistic and dependent on individuating features (e.g., personality or environment; Berndt & Heller, 1986; Biernat, 1991; Taylor, 1996; Taylor et al., 2009). Thus, examining naïve theories across childhood is critical to understanding the developmental course of social cognition.

### *Naïve Sociology*

The aim of the present research is to examine the development of a component of naïve sociology—children's abstract expectations about how group memberships constrain behavior. In describing naïve sociology, Hirschfeld (1996) proposed that an expectation that people can be categorized into discrete kinds—reflecting fundamental similarities among category members and differences between groups—guides children's understanding of the social world. Consistent with this proposal, by preschool, children understand certain social categories as marking stable, fundamental kinds (Hirschfeld, 1995; Rhodes & Gelman, 2009; Taylor, 1996; Taylor et al., 2009), and as indicative of many

---

I am very grateful to the children, parents, and teachers at the preschools and elementary schools that participated in this research, as well as to the families and staff at the Children's Museum of Manhattan. I am also very thankful to Kelli Grobe, Joe Dietzel, Abigail Mengers, Lisa Chalik, and David Berman for their assistance with data collection, to Yarrow Dunham and Amanda Brandone for helpful feedback on an earlier draft of this manuscript, and to Karl Edwards for the study illustrations. Funding was provided by NYU challenge funds and NSF Grant BCS-1147543.

Correspondence concerning this article should be addressed to Marjorie Rhodes, 6 Washington Place, room no. 301, New York, NY 10003. Electronic mail may be sent to marjorie.rhodes@nyu.edu.

© 2012 The Author

Child Development © 2012 Society for Research in Child Development, Inc.

All rights reserved. 0009-3920/2012/8306-0006

DOI: 10.1111/j.1467-8624.2012.01835.x

underlying similarities among category members (e.g., in preferences, skills, and abilities; Diesendruck & haLevi, 2006; Gelman et al., 1986; Kalish & Lawson, 2008; Rhodes & Gelman, 2008; Waxman, 2010). In this manner, children's social categories share representational and inferential features with categories in other domains; for both social categories and animal categories, for example, children treat certain category memberships as stable, fundamental, and inductively informative (Rhodes & Gelman, 2008, 2009; Taylor et al., 2009).

Although category-based induction can operate similarly in the social and biological domains (e.g., in both domains, children base inferences on abstract conceptual categories, instead of on perceptual features or other individuating properties, Diesendruck & haLevi, 2006; Gelman et al., 1986), social categories can also serve inferential roles that are uniquely relevant to the social world. For example, social categories may guide predictions about social interactions, such as whether two people will be friends or enemies, cooperate or compete, or help or harm each other (see Shutts, Roben, & Spelke, in press). Making such predictions is an important component of successfully navigating the social world, yet little prior research has examined whether social categories serve this inferential role in early childhood, or how use of social categories in this manner develops and changes across childhood. Examining this feature of social categories was the goal of the present work.

#### *Older Children's Theories of Groups*

Previous developmental studies examining children's beliefs about how group memberships influence social interactions have primarily focused on the middle and older childhood years (Abrams, Rutland, & Cameron, 2003; Abrams & Rutland, 2008; Abrams, Rutland, Ferrell, & Pelletier, 2008, 2009). For example, Abrams and colleagues have found that older children (ages 8–10) understand that a member of their own group who is loyal only to their own group will be simultaneously preferred by their own group and disliked by an out-group, whereas a member of their own group who shows mixed loyalty will be relatively disliked by their own group, but preferred by the out-group. In this manner, older children have an abstract understanding that loyalty norms govern in-group and out-group behavior. Abrams and colleagues have found that this understanding relies on advanced social perspective taking and theory of mind abilities, as well as children's own experiences in

groups, and thus does not fully develop until later childhood (e.g., beginning around age 7, and becoming more robust around ages 8–10).

Although Abrams and colleagues (Abrams et al., 2003; Abrams et al., 2008, 2009) provide clear evidence that children's use of social groups to understand social interactions develops in important ways throughout childhood, these findings do not preclude the possibility that younger children have rudimentary naïve theories of social groups, which perhaps feed into these later, more sophisticated understandings. The experimental tasks used in prior work have been designed to be more suitable for older children, and children under the age of 5 have not been included in these previous studies. Simplified tasks designed to assess younger children's basic expectations about how groups constrain social interactions are thus necessary to evaluate whether younger children are beginning to form naïve theories of social groups.

#### *Children's Own Group Memberships*

To consider whether preschool-age children have abstract expectations about how groups constrain social interactions, it is useful to consider how young children feel and behave toward members of their own and other groups (Aboud, 1988, 2003; Bigler & Liben, 2007; Nesdale, 2004). By preschool, children prefer friends who share membership with them in a familiar category (e.g., gender, race, or language; Kinzler, Shutts, DeJesus, & Spelke, 2009; Kircher & Furby, 1971; Maccoby & Jacklin, 1987). Also, Patterson and Bigler (2006) found that assigning preschoolers to novel color groups within their school led children—after several weeks spent in the groups—to prefer members of their own group and to expect members of their own group to obtain more positive outcomes (e.g., to win contests; for similar effects in older children, see Bigler, 1995; Bigler, Jones, & Lobliner, 1997; Nesdale & Flessner, 2001; Sherif, Harvey, White, Hood, & Sherif, 1961). These effects were stronger when the groups were labeled and used in a functional manner in the children's classrooms (e.g., used to organize lines to go to lunch); however, young children are also sensitive to arbitrary group memberships immediately after they are assigned. For example, Dunham, Baron, and Carey (2011) assigned children (age 5) to color groups and found that children immediately favored their in-group members on ratings of how much they liked new children (shown in photographs), sharing of resources, behavioral predictions (e.g., who would do positive

actions and obtain positive outcomes), reciprocity (e.g., who a participating child thought would share with them), and on an implicit bias measure. Although children's group memberships influence their feelings and beliefs on a range of experimental tasks, it is less clear that they influence their actual behavior. For example, although Dunham et al. (2011) found that young children said they would share more resources with their in-group members, Patterson and Bigler (2006) found that preschoolers were no more likely to play with members of their own color groups, and Bigler et al. (1997) found that children were no more likely to provide help to an in-group member than an out-group member.

Although children have an early emerging sensitivity to their own group memberships, however, these previous studies do not speak directly to whether children have abstract expectations about how groups constrains social behavior for two reasons. First, the measures in these previous studies have generally included assessments of whether children prefer their own group members, not of whether children use group memberships to predict social interactions (e.g., whether people will help or harm each other). Whether social categories serve this inferential role in early childhood has not been the focus of prior work in this area. Second, when measures have been included that could speak to children's beliefs about social interactions, it has been difficult to determine whether children's responses reflect abstract expectations about how groups constrain social behavior, or alternately, whether such responses reflect more generalized affective processes.

For example, Dunham et al. (2011) asked children to make predictions about reciprocity (e.g., whether an in-group or out-group member would be more likely to share with them) and found that children favored in-group predictions on these items. These findings could reflect an abstract expectation that the members of the same group will share with each other but could also be driven by generalized positivity (e.g., matching the positive behavior *sharing* with generalized positive feelings toward the in-group).

As described by Dunham et al. (2011), three features of their data support the interpretation that generalized positive feelings underlie children's responses across a range of measures. First, children with stronger implicit in-group preferences made more in-group-favoring behavioral predictions (e.g., were more likely to predict that an in-group member would do something nice), suggest-

ing that such predictions relate to general positive affect. Second, children reliably favored in-group members for positive actions but chose in-group members and out-group members equally often for negative actions. This asymmetry is consistent with asymmetries in children's affective biases, as the early effects of group membership often include positive feelings toward in-groups but not necessarily negative feelings toward out-groups (Aboud, 2003; Bigler et al., 1997; Brewer, 1999; Dunham et al., 2011; Nesdale, 2004). Third, and most directly, children also favored their in-group for positive events brought about by chance (e.g., they predicted that their in-group members would be more likely to find a dollar on the street). Together, these data suggest that children's responses may not reflect abstract causal theories about how group memberships influence behavior, but rather more generalized affective processes.

### *The Present Studies*

Thus, previous developmental research has found (a) that categories play an important inductive role in early social cognition (Diesendruck & haLevi, 2006; Gelman et al., 1986), (b) that children's theories of how social groups influence behavior undergo important developments in middle and older childhood (Abrams et al., 2003; Abrams et al., 2008, 2009), and (c) that young children are sensitive to their own group memberships (Dunham et al., 2011; Nesdale, 2004; Patterson & Bigler, 2006). Yet, whether preschool-age children use social categories to make predictions about social interactions—who will be friendly or mean to each other, or who will help or harm each other—has not been examined in prior work and was the focus of the present studies.

One possibility is that children derive abstract expectations about how groups influence social behavior from the feelings and experiences that they themselves have as group members. If so, children's theories of groups should show a similar developmental trajectory as their affective and behavioral responses to their own group memberships. For example, children should use social groups to predict positive behaviors (e.g., that agents will preferentially help members of their own groups) before they use groups to predict negative behaviors (e.g., that agents will preferentially harm members of different groups), just as positive feelings toward in-groups develop prior to negative feelings toward out-groups (Bigler et al., 1997; Dunham et al., 2011; Nesdale, 2004). Yet, another

possibility is that two components of group psychology develop in parallel in early childhood: (a) behavioral and affective biases in favor of in-groups (as documented by Dunham et al., 2011; Patterson & Bigler, 2006) and (b) abstract conceptual expectations of how groups constrain social interactions. If so, then these phenomena may show distinct developmental trajectories. Based on prior work, whether preschoolers have naïve theories of social groups, the types of inferences that such theories support and how these theories change across development are open questions. Thus, these issues were the focus of the present research.

The present studies directly examine inferences about social interactions. Instead of examining predictions about who will do positive or negative actions or obtain positive or negative outcomes (Who will share a cookie? Who will steal a cookie?; Dunham et al., 2011; Patterson & Bigler, 2006), the present studies ask children to predict whom an agent will direct positive or negative actions toward (With whom will someone share a cookie? From whom will someone steal a cookie?). Also, to establish that such predictions reflect abstract expectations about how group members relate to one another, these studies ask children to make predictions involving groups in which children themselves do not hold membership. Thus, children's generalized positive affect toward their own group members cannot underlie predictions in these studies.

These studies introduce children to novel groups ("Flurps" and "Zazes") of which they are not members, and ask them to predict the recipient of an action performed by an agent from one group (e.g., "Here is a Flurp. One day, this Flurp shared a cookie with someone. Who did the Flurp share a cookie with? Did the Flurp share a cookie with another Flurp? Or did the Flurp share a cookie with a Zaz?") Questions involved both helpful behaviors (e.g., sharing) and harmful behaviors (e.g., stealing). If this work had used groups based on familiar categories (e.g., gender), then children could have responded based on prior beliefs about specific group-based conflict (e.g., conflict between boys and girls). Thus, novel categories allow for a clearer test of children's generalized expectations.

In all studies, the groups were marked by shirt color (red shirt or blue shirt) and labeled with a novel noun label ("Flurps" and "Zazes"). Although there is evidence that simple perceptual distinctions and group labels are sufficient to elicit children's own group-based biases (Dunham et al., 2011), previous work has also found that children

attribute more meaning to novel categories that have a functional role in the environment (Bigler et al., 1997; Patterson & Bigler, 2006), particularly when groups are engaged in competition (Rhodes & Brickman, 2011; Spielman, 2000). Because the present study examined different inferences than have been studied in prior work, it was an open question whether the novel groups would need to have these functional characteristics for children to incorporate group membership into their predictions. For example, perhaps children only expect group memberships to influence helpful and harmful social interactions when the groups have a functional role (e.g., when the members of a group are cooperating with each other) or are placed in a competitive context. Alternately, children could rely on naïve theories of social groups more robustly. To address this issue, Study 1 varied the functional information that was provided about the novel groups.

Study 1 examined whether preschool-age children (ages 3–5) use novel social groups to make predictions about social interactions, and tested the types of groups that elicit these predictions. Studies 2 and 3 examined the expectations underlying children's inferences in more detail. Study 4 examined how children's predictions change across childhood (ages 5–10).

## Study 1

Study 1 tests whether preschool-age children use novel social groups to make predictions about helpful and harmful social interactions. Study 1 also examines whether these inferences vary depending on the functionality of the novel groups.

### Method

*Participants.* Participants included 96 preschoolers (48 boys, 48 girls; 72% White, 11% Multiethnic, 4% Asian or Asian American, 6% Hispanic, 6% did not report ethnicity) recruited from preschools across the borough of Manhattan in New York City. Of the 96 preschoolers, 49 were 3-year-olds ( $M$  age = 3.43 years, range = 3.0–3.99), and 47 were 4- and 5-year-olds ( $M$  age = 4.75 years, range = 4.05–5.99). Parents provided written consent and completed a family demographic form reporting the child's gender, birthdate, ethnicity, the languages spoken at home and by the child, the number of children at home, and the participating child's place in the birth order.

*Procedure.* Children participated in a quiet area of their school. A trained research assistant administered the experiment individually. First, children were introduced to the novel categories of people. To evaluate whether children's inferences depend on the functional properties of the novel groups, children were randomly assigned to one of three ways of introducing the novel categories. These included: (a) minimal (the groups were labeled and perceptually marked only), (b) no competition (the groups were labeled, perceptually marked, and described as having a functional role in the agents' environment, but there was no between-group competition), and (c) competition (the groups were labeled, perceptually marked, described as having a functional role, and were presented in the context of between-group competition). After introduction to the stories (see next), children completed predictions about either helpful or harmful behaviors. Thus, children were randomly assigned to condition following a 3 (group context: competition, no competition, and minimal)  $\times$  2 (behavior: harmful and helpful) factorial design. There were 16 children in each condition.

*Introduction to the novel groups.* First, children in all conditions were introduced to the novel groups by showing them pictures of each group, and stating, for example, "Here is the red team, they are called the Zazes. Here is the blue team, they are called the Flurps." In the minimal context, children received no other information about the groups. In both the competition and no competition contexts, children were then read a story that described each group working to build a block tower. The no competition context described within-group cooperation, but no between-group competition (there were enough blocks for both towers, and both groups would receive prizes in the end). The competition context described both within-group cooperation and between-group competition (there were not enough blocks for both towers, and only one group would receive a prize in the end). None of the introductory stories described any interactions between members of different groups. The pictures shown to children and the full text of the introductory stories are presented in Appendix S1 (online supporting information).

*Comprehension questions.* During the stories, children were asked two comprehension questions, including (a) to point to the Flurps and Zazes, and (b) whether there were enough blocks for each group. If children missed a comprehension question, the story was repeated, and the questions were re-asked. No child failed the comprehension questions a second time.

*Test questions.* Following the introduction to the groups, children completed six test questions. For each question, children were told that an agent performed an action, and their task was to predict the recipient. For example, children heard, "One day a Flurp hit someone. Who did the Flurp hit? Did the Flurp hit another Flurp? Or did the Flurp hit a Zaz?" Children were shown two pictures, one depicting a Flurp hitting another Flurp, and one depicting a Flurp hitting a Zaz. Children responded by pointing to the picture that depicted their expectation. For both helpful and harmful behaviors, three of the six questions involved the setting described in the stories (block building) and three involved a new setting (a playground). A sample test question is presented in Appendix S1; the content of all of the test questions is summarized in Table 1.

*Counter balancing.* The following factors were counter balanced across participants: whether the red group was referred to as "Flurps" or "Zazes," whether the red group was shown on the participant's left or right, which group was mentioned first in the story, the group membership of the agent in the test questions, the lateral position of the two answer choices, and whether the test questions referring to the block or playground setting were asked first.

*Scoring.* For each question, children could choose the image that showed the agent directing the behavior toward a member of the agent's group (e.g., a Flurp hugging another Flurp; referred to as a within-group prediction; scored as 0) or the image that showed the agent directing the behavior toward a member of the other group (e.g., a Flurp hugging a Zaz; referred to as a between-group prediction; scored as 1). Based on this scoring system, if children expect agents to harm members of different groups but to help members of their own, we should find high scores for harmful behaviors (i.e., more predictions of between-group harm), and low scores for helpful behaviors (i.e., fewer predictions of between-group helping, and thus more predictions of within-group helping).

*Justifications.* After the last test question, children were asked to explain their response. For example, after children responded to the question, "Who did a Flurp say could not play with him?" children were asked, "Why did the Flurp say the [Flurp or Zaz] could not play?" Children's responses were recorded verbatim and subsequently checked for accuracy against videotapes of the testing sessions. Two independent raters coded each response. Explanations were coded as individual, if they

Table 1  
Summary of Test Questions for Harmful Behaviors and Helpful Behaviors

Harmful	Helpful
Block setting	
Who did a Flurp steal a block from?	Who did a Flurp share a block with?
Who did a Flurp hit while building a block tower?	Who did a Flurp hug while building a block tower?
Who did a Flurp say could not help build a block tower?	Who did the Flurp say could help build a block tower?
Playground setting	
Who did a Flurp steal a cookie from?	Who did a Flurp share a cookie with?
Who did a Flurp hit on the playground?	Who did a Flurp hug on the playground?
Who did a Flurp say he would not play with on the playground?	Who did a Flurp play with on the playground?

referred to an individual property of the agent (e.g., "he is mad," "he is mean," "he does not like him") and as category-based if they referred to group membership (e.g., "because they are in different groups," "because he is a Flurp but he is a Zaz.") Interrater agreement was .89, with disagreements resolved by discussion.

### Results

*Preliminary analyses.* Preliminary analyses revealed no effects of any of the variables relating to story or test-question counter-balancing (here, or in subsequent studies). Preliminary analyses also revealed no effects of participant gender, whether the child had siblings, whether the child was exposed to one or multiple languages at home (here, or in subsequent studies). None of these variables were considered further.

*Analysis plan.* Data were composed of a series of responses to binomial trials (i.e., children could receive a 1 or a 0 for each of six trials). To respect the dichotomous nature of the dependent measure, data were analyzed with binomial regression models. Initially, a binomial regression model tested the main and interactive effects of three factors: group context (no competition, competition, and minimal), behavior (helpful and harmful), and test question setting (playground and block). This analysis revealed effects of group context and behavior (see next), but no main or interactive effects of test question setting (here, or in subsequent studies),  $ps > .50$ . Indeed, inspection of the means revealed identical effects for responses to these items (see Appendix S2). Therefore, this factor was not considered further.

In the main analyses, group context (competition, no competition, and minimal) and behavior (helpful and harmful) were entered as fixed factors, and analyses tested for a main effect of each factor,

as well as for their interaction. The dependent variable is the number of times that children made between-group predictions, of the total possible (6). These analyses yield Wald  $\chi^2$  values as indicators of significant effects. For ease of interpretation, data are presented as probabilities (the probability of a between-group prediction), accompanied by Wald 95% confidence intervals (CIs). When these CIs do not include .50 (the probability expected by chance), this indicates that pattern deviated from equal probability responding. To facilitate comparison with previous work, Cohen's  $d$  values are reported as indicators of effect sizes. All post hoc comparisons report  $p$  values adjusted with sequential Bonferroni corrections.

*Main analyses.* The binomial regression model revealed significant main effects of group context,  $\chi^2(2) = 9.62$ ,  $p = .008$ ; behavior,  $\chi^2(1) = 22.50$ ,  $p < .001$ ; and an interaction between group context and behavior,  $\chi^2(2) = 12.76$ ,  $p = .002$ . The main effect of behavior operated such that children made more between-group predictions for harmful behaviors ( $M = .77$ ,  $CI = .72, .82$ ) than for helpful behaviors ( $M = .58$ ,  $CI = .52, .64$ ). The main effect of group context operated such that children made more between-group predictions in the no competition ( $M = .72$ ,  $CI = .65, .79$ ) and competition ( $M = .73$ ,  $CI = .66, .80$ ) contexts than in the minimal context ( $M = .59$ ,  $CI = .52, .66$ ),  $ps < .01$ .

The interaction between group context and behavior is presented in Figure 1. In both the no competition and competition contexts, children made more between-group predictions for harmful behaviors than for helpful behaviors ( $ps = .001$ , no competition,  $d = 1.0$ ; competition,  $d = .82$ ), but they did not do so in the minimal context. Also, as can be seen from the 95% CIs (see Figure 1), children in the competition and no competition contexts made between-group predictions for harmful behaviors more often than expected by chance,  $ps < .001$ , but

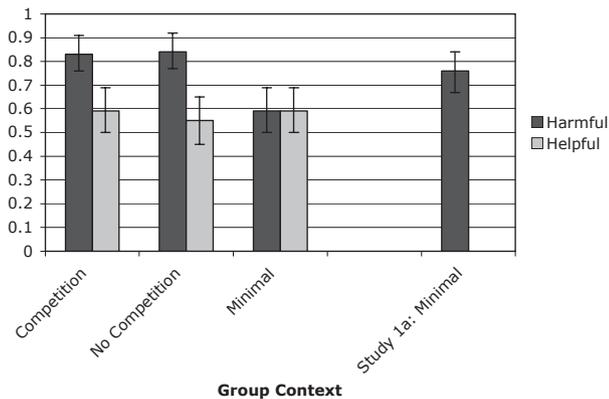


Figure 1. Probabilities of between-group predictions for Studies 1 and 1a, by group context and behavior type, with Wald 95% confidence intervals.

their responses for helpful behaviors did not differ from chance. In the minimal context, children's responses did not differ from chance for either type of behavior. There were no differences between the no competition and the competition contexts.

*Justifications.* For harmful actions, chi-square analyses revealed that children's explanations varied by group context,  $\chi^2(4) = 11.70$ ,  $p = .02$ . Category-based explanations were more prevalent in the no competition context (six category-based, seven individual, three no answer) and the competition context (six category-based, one individual, nine no answer) than in the minimal context (one category-based, six individual, nine no answer). Thus, children in the no competition and competition contexts used group membership to both predict and explain harmful actions. For helpful behaviors, chi-squares analyses revealed that children's explanations did not vary by group context,  $\chi^2(4) = 4.93$ ,  $p > .30$ . For helpful behaviors, children gave 10 category-based explanations, 14 individual explanations, and 24 no responses.

*Age.* There were not sufficient numbers of 3-, 4-, and 5-year-olds to fully examine the effects of age within the preschool period. It is useful to consider, however, whether the pattern of effects was found when the youngest participants (3-year-olds;  $N = 49$ , evenly divided across conditions) are examined alone. Indeed, 3-year-olds predicted between-group harm more often than expected by chance in the no competition context ( $M = .86$ ,  $CI = .75, .97$ ,  $p < .001$ ) and the competition context ( $M = .72$ ,  $CI = .58, .87$ ,  $p = .01$ ), but they did not do so in the minimal context ( $M = .55$ ,  $CI = .44, .65$ ). For helpful behaviors, their predictions did not differ from chance in any group context, all  $ps > .30$ . Thus, younger children showed the same pattern as the sample as a whole.

## Discussion

In Study 1, preschoolers reliably predicted between-group harm, but they predicted between-group helping and within-group helping equally often. These data suggest that preschool-age children have a naïve theory of social groups that guides their inferences about harm—they expect agents to harm members of other groups rather than members of their own. Yet, preschoolers' naïve theories of social groups do not appear to influence inferences about helping.

Children's inferences in Study 1 do not reflect a simple extension of observed information. This study presented novel groups; thus, children did not enter the experiment with previous knowledge or beliefs about the social history between the groups. Also, no interactions between members of different groups—or negative interactions of any kind—were described in the introductory stories. If children based their inferences on directly observed information, they should have reliably predicted within-group helpful actions (as only within-group positive behaviors were shown in the story), or even reliably predicted within-group interactions for both types of behaviors (as the characters in the story interacted only with members of their own groups). Yet, children's inferences did not fit any of these patterns; children reliably predicted between-group harm, a type of interaction not shown in the story.

Study 1 suggests that children have abstract, naïve theories of social groups that underlie their inferences about harm. One way that children could have responded to these questions without relying on abstract theories would have been to identify with the agent in the test questions, to develop positive affect toward the agents' group members (as if they were the child's own group members), and then to base responses on those affective biases. The obtained pattern, however, is inconsistent with this possibility. Such a process should have led to reliable predictions that in-group members would receive the helpful actions (as children expect their in-group members to obtain positive outcomes; Dunham et al., 2011; Patterson & Bigler, 2006), but should not have led to predictions that out-group members would receive the harmful actions (as children predict that in-group and out-group members are just as likely to obtain negative outcomes; Dunham et al., 2011). Thus, the obtained effects show the reverse pattern as would have been found if identification with the agent—and resulting affective biases—drove children's responses.

Study 1 speaks to the types of categories that elicit children's predictions of between-group harm, yet leaves some issues unresolved. In Study 1, children reliably predicted between-group harm in both the competition and no competition contexts, but they did not do so in the minimal context. The data from the competition and the no competition conditions suggest that intergroup competition is not necessary for preschoolers to expect between-group harm. It is possible, however, that because the groups were always referred to as "teams," children in the no competition context inferred that the groups would compete, even though competition was not mentioned in the story (and, in fact, the story emphasized the noncompetitiveness of the situation). Also, the null effects in the minimal condition may indicate that functional information about the groups (e.g., that group members cooperate with each other to achieve group-relevant goals; Bigler & Liben, 2007) may be necessary to elicit preschoolers' naïve theories. Yet, children in the minimal context also had less time to process the novel groups (as the introductory story was briefer). Thus, either differences in the amount of exposure to groups or differences in the functional significance of groups could account for the difference between the minimal context and the other two contexts. Study 1a was designed to address these issues.

### Study 1a

The aim of Study 1a was to test whether children would use novel social groups to make predictions about harmful actions in the absence of functional information, when the novel categories were made more salient and children were given longer to learn them (in comparison to the Minimal context conditions of Study 1). This study also tests whether children will predict between-group harm when the groups are referred to simply as "groups" instead of as "teams."

*Method.* Participants included 14 children recruited from the Children's Museum of Manhattan (7 boys, 7 girls,  $M$  age = 3.67 years, range = 3.01–4.73 years). Race or ethnicity information was not available for the sample of children who participated in this particular study, but overall, the sample of families who participate in research through this site is approximately 65% White, 10% African American, 10% Asian American, 7% Hispanic, and 8% Other.

Children were introduced to the novel groups, without receiving any functional information. Children were simply shown four individual group members from each group (see top panel,

Appendix S1) and told, "This is the blue group; they are called the Flurps. This is the red group; they are called the Zazes. Can you point the Flurps? Can you point the Zazes?" To increase the salience of the groups, the experimenter repeated this procedure (the experimenter pointed to and labeled each group a second time and then asked the child to point to the groups again). The experimenter then pointed to a character from one of the groups and asked, "Tell me, is this one a Flurp or a Zaz?" and then repeated this question for a character from the other group. The experimenter then repeated the labels a final time, stating, "That is right. These are the Flurps and these are the Zazes." After the introduction to the groups, children were asked to make predictions about the same six harmful actions asked about in Study 1.

*Results and discussion.* In Study 1a, children predicted between-group harm more often than expected by chance, Wald  $\chi^2(1) = 23.33$ ,  $p < .001$  ( $M = .76$ ,  $CI = .67, .84$ ). These data demonstrate that when novel groups are made more salient (through increased exposure time and repetition of the category labels), children use novel social groups to predict between-group harm, even in the absence of functional information (and even when the groups are referred to as "groups" instead of as "teams"). Comparing children's predictions for harm in Study 1a to predictions for harm in each of the three group contexts from Study 1 revealed a main effect of group context, Wald  $\chi^2(3) = 19.85$ ,  $p < .001$ . The new minimal condition did not differ from the competition or no competition conditions of Study 1,  $ps > .40$ , but did differ from the minimal condition of Study 1,  $p = .05$  (see Figure 1). Study 1a shows that when novel minimal groups are made sufficiently salient, children invoke their naïve theories to make predictions about harm. Thus, neither functional information nor a competitive context is necessary to elicit children's predictions of between-group harm. The data from the competition and no competition conditions of Study 1, along with the data from Study 1a, show that preschool-age children robustly predict between-group harm, across multiple ways of defining novel social groups. The aim of Studies 2–3 was to examine the expectations underlying children's predictions of between-group harm in more detail.

### Study 2

Study 1 found that preschool-age children reliably predict between-group harm. There are two independent expectations that could underlie these

predictions: (a) an expectation that individuals will *not* harm members of their own group and (b) an expectation that individuals *will* harm members of another group. The aim of Study 2 was to examine whether preschoolers have each of these independent expectations, toward the goal of better characterizing whether their naïve theories emphasize expectations about how members of the same group relate to one another (e.g., that agents will refrain from harming their own group members), expectations about how members of different groups relate to one another (e.g., that agents will direct harm toward members of different groups), or both.

### Method

Participants included 32 preschoolers recruited from the Children's Museum of Manhattan (10 male, 22 female;  $M$  age = 3.7 years, range = 3.0–4.9). Procedures were identical to Study 1, except as follows. In this study, all participants heard the no competition story (see Appendix S1). Then they were told, "I'm going to show you some pictures of children and ask you to make some guesses about things they did. Some of them will be [Flurps or Zazes] like the children in the story that we just read, and some of them will just be kids who are not Flurps or Zazes at all. The kids who are not Flurps or Zazes did not play with the blocks at all."

Children were assigned to one of two conditions. In the within-group versus no-group condition, test questions presented a choice between within-group harm (scored 0) and harm directed toward the character who was not a member of either group (scored 1). For example, "One day a Flurp hit somebody. Who did the Flurp hit? Did the Flurp hit another Flurp? Or did the Flurp hit this kid who is not a Flurp?" In the between-group versus no-group condition, test questions presented a choice between between-group harm (scored 1) and harm directed toward the character who was not a member of either group (scored 0). For example, "One day a Flurp hit somebody. Who did the Flurp hit? Did the Flurp hit a Zaz? Or did the Flurp hit this kid who is not a Zaz?" The test questions were presented in the same visual format as shown in Appendix S1, with the exception that the "kid who was not a Flurp or Zaz" wore a shirt that was not blue or red. The color shirt on these characters varied across test questions to emphasize that these characters were not members of any group (e.g., on one item, the child wore a gray shirt, on another a yellow shirt, and so on).

All children completed six test questions about the same types of harmful actions that were asked about in Study 1 (see Table 1). Because children may have been biased against selecting the "no-group" recipient for test questions that referred to blocks (since blocks were previously associated only with Flurps and Zazes), however, analyses focused only on the test questions that involved events that occurred on a playground. The same pattern was also found when all the test questions were considered together.

### Results and Discussion

In the within-group versus no-group condition, preschoolers reliably predicted that agents would harm people who are not members of their group ( $M = .78$ ,  $CI = .65, .87$ ),  $p < .001$ . In the between-group versus no-group condition, preschoolers reliably predicted that agents would harm members of the contrasting group ( $M = .67$ ,  $CI = .51, .79$ ),  $p = .03$ . Thus, preschoolers expect that agents will harm members of contrasting groups over people with no salient group membership, but will harm people with no salient group membership over members of their own group. This pattern suggests that preschoolers' naïve theories include both expectations of how people interact with their own groups (refrain from harm) and how people act toward contrasting groups (engage in harm).

Across the two conditions in this experiment, a pattern in which children simply selected the more salient recipient cannot explain the findings. In the within-group versus no-group condition, for example, the recipient in the within-group response option (e.g., a Flurp hitting a Flurp) should have been more salient, as this character was from a group mentioned in the introductory story, and the other recipient was not. Yet, children favored the other response choice on these items (e.g., they selected responses where a Flurp hit a "kid who was not a Flurp or a Zaz"). Because this condition is not consistent with a salience-based account, these data provide strong evidence that children expect agents not to harm members of their own groups. In the between-group versus no-group condition, however, the data are consistent with the possibility that children selected the response option where the recipient was simply more salient (because the favored recipient was from a group mentioned in the story). Thus, the conclusion that children reliably expect agents to harm members of contrasting groups is more tentative, and should be examined in more detail in future work.

### Study 3

In Studies 1–2, children reliably predicted between-group harm, suggesting that they have a naïve theory of social groups that guides their expectations about social interactions. An alternate possibility, however, is that children have used a simpler heuristic, namely, that they expect negative outcomes to happen to an agent's out-group members, without necessarily considering the role of social interactions. Thus, perhaps they expect that a negative action (e.g., having one's cookie stolen) will happen to the agent's out-group instead of the agent's in-group, without considering the agent's role in determining that outcome (e.g., that an agent would cause the cookie to be stolen from the agent's out-group instead of the agent's in-group). The aim of Study 3 was to address this possibility.

#### Method

Participants included 16 preschoolers recruited from the Children's Museum of Manhattan (10 boys, 6 girls;  $M = 3.78$  years, range = 3.0–4.55 years). All procedures were identical to Study 1, except as follows. All participants heard the no-competition story (see Appendix S1). Then, participants were asked a series of 12 questions, including 6 helpful actions and 6 harmful actions, in counter balanced order. The questions asked about similar content as those in Study 1 (see Table 1). Unlike in previous studies, however, these questions asked children to predict whom the agent would see as the recipient of each action (e.g., "One day, a Flurp saw somebody's cookie get stolen. Whose cookie was stolen? Was the Flurp's cookie stolen? Or was the Zaz's cookie stolen?"). For these items, a picture of the agent (the person who saw the event) was provided, to help children remember who saw the event take place. The two response options (e.g., a Flurp's cookie being stolen and a Zaz's cookie being stolen) were placed equidistant from the agent. There was no agent shown in the pictures that corresponded to the response options. These questions ask who would receive positive or negative actions but provide no information about who performed the actions. Thus, if children simply expect negative things to happen to members of the agent's out-group, then, as in prior studies, they should reliably select the recipient who is a member of the agent's out-group. In contrast, if their performance in previous studies reflected systematic expectations about social interactions, they should perform at chance in this study.

#### Results and Discussion

Preschoolers' predictions did not differ from chance for either harmful behaviors ( $M = .52$ ,  $CI = .42, .63$ ) or helpful behaviors ( $M = .49$ ,  $CI = .39, .59$ ),  $ps > .50$ . There were no effects of order,  $ps > .80$ ; children responded at chance for harmful actions regardless of whether they were asked about first ( $M = .54$ ,  $CI = .40, .68$ ) or second ( $M = .50$ ,  $CI = .35, .65$ ) and for helpful actions regardless of whether they were asked about first ( $M = .48$ ,  $CI = .33, .63$ ) or second ( $M = .50$ ,  $CI = .36, .64$ ).

These data suggest that preschoolers do not have simple expectations that good things will happen to an agent's in-group or that bad things will happen to an agent's out-group (or that agents will see good things happen to their in-group, but bad things happen to their out-group). One way that children could have responded to these questions would have been to infer that agents are more likely to see things happen to their in-group members *in general*. Yet, children did not reliably select the agent's in-group members on these items, for either helpful or harmful behaviors, suggesting that they did not rely on this type of baseline assumption in this context. Considering the data from Study 3 and those from Studies 1–2 together suggests that preschool-age children reliably expect harmful things to happen to an agent's out-group only in the context of social interactions. That is, they expect the agent to *direct* harmful actions toward a member of the agent's out-group. Thus, Study 3 confirms that children's responses in Studies 1–2 reflect their naïve theories of how groups constrain social interactions.

### Study 4

Studies 1–2 found that by age 3, children have a generalized expectation that agents will direct harm toward different groups, but not that agents direct help toward their own groups. Perhaps beliefs about the role of groups in determining helpful behaviors depends on more extended developmental processes, such as the development of social perspective taking, theory of mind, or more time spent in groups (Abrams et al., 2008, 2009). If so, identifying the age at which children begin to reliably predict within-group helping may be informative regarding the underlying process. Thus, Study 4 examined the developmental trajectory of these predictions.

### Method

**Participants.** Participants included 75 children ranging in age from 5 to 10 (60% White, 4% Hispanic, 3% Asian or Asian American, 13% reported more than one racial or ethnic background; the remainder chose not to report race or ethnicity). These included 15 kindergarteners (8 boys, 7 girls;  $M$  age = 5.71 years, range = 5.30–6.12), 19 first graders (9 boys, 10 girls,  $M$  age = 6.64 years, range = 6.18–7.10), 20 second graders (12 boys, 8 girls,  $M$  age = 7.62 years, range = 7.23–8.44), and 21 third and fourth graders (10 boys, 11 girls,  $M$  age = 9.31, range = 8.30–10.23). All children were recruited from a single public elementary school in New York City and were tested on-site at their elementary school in a quiet area.

**Procedures.** All procedures were identical to Study 1, except as follows. First, all children in Study 4 were told the no competition story (see Appendix S1). Second, all children completed test items about helpful and harmful behaviors (see Table 1), in counter balanced order across participants.

Children were asked to explain their prediction for the last question in each block (following the procedures of Study 1). Because this study involved older children, children gave more elaborate justifications that allowed for a more detailed coding scheme. Responses were coded as *agent's own feelings* if the child referred to the agent's individual motivation (e.g., "he liked her," "he wanted to be nice," "he was angry"), *relationships* if the child referred to the relationship between the agent and recipient (e.g., "they are friends," "they are enemies"), *rule following* if they referred to conformity to a social rule (e.g., "it was a rule that he had to share with him"), or *groups* if they referred to group membership (e.g., "because they are both Flurps," "Because they are in different groups"). Inter-rater reliability was .85 with discrepancies resolved through discussion.

### Results

Data were analyzed through a 2 (behavior: helpful and harmful)  $\times$  4 (grade level) binomial regression model, with behavior as a within-subjects factor. This analysis revealed a main effect of behavior,  $\chi^2(1) = 45.30$ ,  $p < .001$ , and a marginal effect of grade level,  $\chi^2(3) = 6.61$ ,  $p = .085$ . Children made more between-group predictions for harmful behaviors ( $M = .88$ ,  $CI = .82, .94$ ) than for helpful behaviors ( $M = .34$ ,  $CI = .25, .44$ ),  $d = 1.61$ . Like

the preschoolers in Study 1, children reliably predicted between-group harm,  $p < .001$ . Children in Study 4, however, also reliably predicted within-group helping (they favored between-group predictions for helping less often than expected by chance,  $p < .01$ ). The marginal effect of grade level indicated that kindergarteners made more between-group predictions than children of any other grade,  $ps < .05$  ( $M_{\text{kindergarten}} = .77$ ,  $CI = .67, .88$ ;  $M_{\text{first grade}} = .57$ ,  $CI = .43, .71$ ;  $M_{\text{second grade}} = .58$ ,  $CI = .44, .73$ ;  $M_{\text{third and fourth grades}} = .68$ ,  $CI = .53, .83$ ).

Although the interaction between grade level and behavior type did not reach significance,  $\chi^2(3) = 4.87$ ,  $p = .18$ , examination of the mean probabilities for helpful and harmful behaviors separately by age indicated that the youngest children—kindergarteners—had a different pattern than that of children from each other group (see Figure 2). Indeed, collapsing the three older grades, and comparing this older age group to the youngest children revealed main effects of behavior,  $\chi^2(1) = 29.08$ ,  $p < .001$ ; age group,  $\chi^2(1) = 5.45$ ,  $p = .02$ ; and a Behavior  $\times$  Age Group interaction,  $\chi^2(1) = 4.68$ ,  $p = .03$ . For harmful behaviors, there was no effect of age group; older children ( $M = .87$ ,  $CI = .73, .94$ ) and younger children ( $M = .87$ ,  $CI = .79, .92$ ) all reliably predicted between-group harm,  $ps < .001$ . For helpful behaviors, however, there was a significant effect of age group; older children made between-group predictions for helpful behaviors ( $M = .26$ ,  $CI = .18, .36$ ) less often than younger children did ( $M = .64$ ,  $CI = .43, .81$ ). Older children reliably favored within-group predictions for helpful behaviors,  $p < .001$ , whereas younger

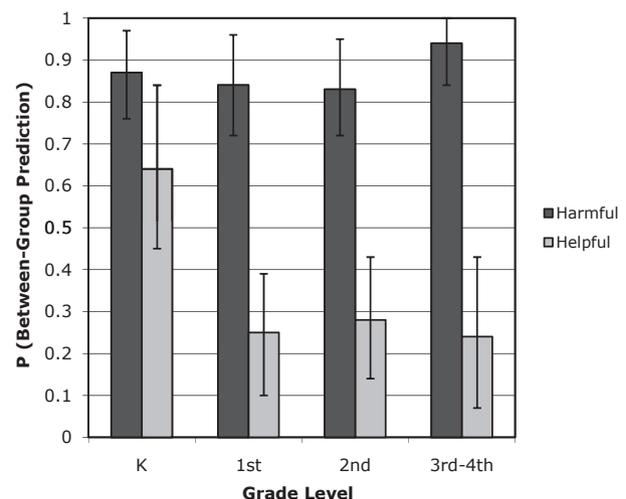


Figure 2. Probabilities of between-group predictions for Study 4, by grade level and behavior type, with Wald 95% confidence intervals.

children’s responses did not differ from chance on these items.

To confirm whether children of each age reliably predicted between-group harm and within-group helping, binomial regression models tested for the effect of behavior type separately for each grade level (see Figure 2). The effect of behavior was marginal for kindergarteners,  $\chi^2(1) = 3.57, p = .059, d = .67$ . As shown in Figure 2, kindergarteners reliably predicted between-group harm,  $p < .001$ , but their predictions for helping did not differ from chance. There were significant effects of behavior for first graders,  $\chi^2(1) = 18.86, p < .001, d = 1.94$ ; second graders,  $\chi^2(1) = 23.29, p < .001, d = 1.76$ ; and third to fourth graders,  $\chi^2(1) = 11.69, p = .001, d = 1.27$ . As shown in Figure 2, children of each of these ages reliably predicted both between-group harm and within-group helping. Thus, expectations of between-group harm remain consistent across childhood, and expectations of within-group helping develop by ages 6–7.

*Justifications.* The most common explanations were group (60% help, 71% harm), followed by relationships (15% help, 7% harm), agent’s own feelings (9% harm, 4% help), and rule following (4% help, 1% harm); the remainder did not fit these codes or did not respond (13% help, 17% harm). For harm, explanations did not differ across grade level; group explanations were the most common justification given at each age (see Table 2). For helpful behaviors, explanations differed marginally across grade level,  $\chi^2(12) = 19.56, p = .076$ . As shown in Table 2, kindergarteners appeared to give a different pattern than any other age group. Indeed, collapsing the three groups of older children and comparing this group with kindergarteners indicated that the pattern of explanations significantly changed across age,  $\chi^2(4) = 15.60, p = .004$ . As shown in Table 2, references to group membership increased with age for helping behaviors.

*Discussion*

In Study 4, kindergarteners, like the preschoolers in Study 1, reliably predicted between-group harm, but responded at chance for helpful actions. By first grade, children reliably predicted both between-group harm and within-group helping. Developmental changes in beliefs about helping were evident in both predictions and explanations. Interestingly, reliable predictions of within-group helping emerged by ages 6–7, which is approximately the same age as Abrams et al. (2008, 2009) have found that children begin to understand how

Table 2  
Numbers of Each Type of Explanation Given by Children of Each Age for Harmful and Helpful Behaviors, Study 4

	Type of explanation				
	Groups	Relationships	Agent’s feelings	Rules	Other
<b>Harm</b>					
Kindergarten	9	2	1	0	3
First grade	16	1	0	1	1
Second grade	12	0	1	0	7
Third to fourth grades	16	2	1	0	2
<b>Help</b>					
Kindergarten	4	4	3	2	2
First grade	12	3	0	0	4
Second grade	13	3	2	0	2
Third to fourth grades	16	1	2	0	2

abstract loyalty norms govern in-group and out-group behavior, although this understanding continues to become more robust in older childhood. Thus, similar developmental processes as those identified by Abrams et al. (2009)—in particular, more advanced theory of mind abilities—may underlie the developmental effects found in the present study. Direct examination of such processes is an important area for future work.

**General Discussion**

In these studies, children relied on naïve theories of social groups to make predictions about individual action. Preschoolers reliably predicted that an agent would harm members of the agent’s out-group (they reliably predicted between-group harm), but did not use group membership to make predictions about helpful behaviors (they predicted within-group and between-group helping equally often). Older children (ages: 6–10) reliably predicted both between-group harm and within-group helping. Preschoolers predicted between-group harm across a range of different group contexts, including when the groups were engaged in within-group cooperation and between-group competition, when they were engaged in within-group cooperation but no between-group competition, and when the groups were marked only by perceptual features and category labels. Children in these studies predicted between-group harm even though they had no evidence of prior negative interactions—or interactions of any kind—between members of different

groups. Thus, children's inferences in these studies reflect their generalized, abstract expectations about how groups constrain social interactions.

The present findings contribute to the literatures on naïve sociology (Hirschfeld, 1996; Olson & Dweck, 2008) and on group membership (Bigler & Liben, 2007; Dunham et al., 2011; Nesdale, 2004). Whereas previous studies have found that children expect category members to be fundamentally similar to each other—and thus to share many novel properties (Diesendruck & haLevi, 2006; Gelman et al., 1986)—the present studies suggest that categories also serve an additional inferential role in early social cognition, by guiding children's predictions about social interactions. These questions are qualitatively different from those asked in prior work on social category-based induction; instead of testing inferences about how particular properties characterize groups, these studies tested inferences about how group members relate to one another and to members of other groups.

Children's inferences in these studies also reflect different processes than those examined in previous studies on group membership. Previous research has found that children generalize positive feelings about themselves to other members of their groups, thus producing group biases in their attitudes and beliefs (Bigler et al., 1997; Dunham et al., 2011; Nesdale & Flessner, 2001). If this process drove the present findings—for example, if children identified with the agent in the test questions and developed positive feelings toward the agent's group members—a different pattern would have been found. In particular, such a process would have led children to reliably predict within-group helping (as children expect good things to happen to their in-group members, Dunham et al., 2011; Patterson & Bigler, 2006), but to respond at chance for harm (as children expect bad things to happen to in-group and out-group members equally often; Dunham et al., 2011), as described in the discussion of Study 1. Thus, the present findings appear not to reflect affective processes driven by group membership.

The present studies differed methodologically from prior work in two key ways, which were both intended to allow these questions to assess children's abstract expectations about how groups constrain social interactions, instead of affective biases. First, in the present studies, children were not members of the relevant groups. Second, children were asked to make different behavioral predictions; instead of asking who will do positive or negative actions (Who will share a cookie? Who will steal a cookie? Bigler et al., 1997; Dunham et al.,

2011; Patterson & Bigler, 2006), these questions asked about social interactions (e.g., With whom will an agent share a cookie? From whom will an agent steal a cookie?). Comparing the present findings to prior work indicates that these two types of predictions rely on different underlying mechanisms; responses to the first type on generalized positive affect toward in-group members, responses to the second on abstract expectations about social groups.

These data show that children's affective responses to their own group members and their abstract expectations about how groups influence behavior follow different developmental trajectories. For affective responses to group members, positive feelings toward in-group members develop prior to negative feelings toward out-group members (Dunham et al., 2011; Nesdale, 2004). In contrast, the present studies show that for children's abstract theories of groups, expectations about negative between-group interactions develop prior to expectations about positive within-group interactions. Thus, these findings indicate that two components of group psychology develop in parallel in the early childhood years: (a) affective and behavioral biases in favor of in-groups, and (b) abstract conceptual expectations about how groups constrain social behavior.

A key area for future work is to examine why expectations of between-group harm develop at an earlier age (by age 3) than expectations of within-group helping (by age 6). Given that young children indicate that they will share more resources with their in-groups, for example (Dunham et al., 2011), it may seem surprising that preschool-age children did not predict within-group helping on these items. Two developmental considerations may help to explain this discrepancy. First, by preschool, children may not have had enough group-based experiences for them to induce a general expectation that people in the same group usually help each other. The preschoolers in the present study were younger than the children in Dunham et al. (2011). Also, Patterson and Bigler (2006) found that even though preschool-age children reported in-group favoring preferences, these preferences did not lead to actual preferential treatment of in-group members (see also Bigler et al., 1997). Thus, children of these ages may have had few actual experiences in which helpful behaviors varied by group membership. Second, even if preschool-age children realize that they themselves prefer to help their in-group members, they may lack the requisite social-cognitive abilities to take the perspective of

another agent, and to infer that the agent would prefer to help in-group members, too. Consistent with these proposals, Abrams et al. (2009) report that increased experiences in groups and more advanced social-cognitive skills contribute to the development of more advanced theories of group interactions.

Given these developmental considerations, however, how do children develop reliable expectations of between-group harm by age 3? Three possibilities will be described here, which are not mutually exclusive. First, expectations of between-group harm could reflect children's previous experiences. It seems unlikely that children directly experience more between-group than within-group harm, as young preschoolers generally spend little time in organized groups (Benenson, Antonellis, Cotton, Noddin, & Campbell, 2008; Benenson, Apostoleris, & Parnass, 1997), and even when they do, group membership often influences attitudes and beliefs, but not actual behavior, as described earlier (Bigler et al., 1997; Patterson & Bigler, 2006; but see Katz, Katz, & Cohen, 1976). Regardless of the actual frequency with which children experience between-group harm, however, an interesting possibility to consider in future work is that children expect between-group harm because such experiences are particularly well represented in memory. For example, if preschoolers view between-group harm as more threatening than within-group harm, such experiences may be better represented in memory (Kinzler & Shutts, 2008), leading to exaggerated estimates of their frequency.

Second, expectations of between-group harm could be supported by specialized cognitive mechanisms for representing intergroup conflict, and thus develop even in the absence of direct input or experience. Cosmides, Tooby, and Kurzban (2003) have proposed that people have specialized cognitive mechanisms for tracking patterns of social allegiance and intergroup conflict. Such mechanisms could produce sensitivity to intergroup conflict—or even lead children to view conflict as the default state of group relations—and thus license predictions of between-group harm based on limited evidence. Whereas there is some evidence that adults have specialized mechanisms for tracking social allegiances and intergroup conflict (Cosmides et al., 2003), whether such mechanisms operate in early childhood—and how such mechanisms support predictions of action—remain open questions (Rhodes, in press).

Third, preschoolers' expectations could reflect their beliefs about moral obligations. Preschoolers

generally predict that people will behave in a manner that is consistent with their obligations (Kalish & Shiverick, 2004). Thus, although the present studies examined expectations about how people *will* behave, not beliefs about how people *should* behave, children's inferences may indeed reflect their beliefs about normative obligations. On this account, perhaps children view within-group harm as more morally problematic than between-group harm, and thus predict between-group harm because they view it as less violating of the agent's obligations.

Counter to this possibility, much research has suggested that preschoolers tend to reject actions that they perceive as treating people unfairly because of group memberships (Bigler, Arthur, Hughes, & Patterson, 2008; Killen, 2007; Killen, Pisacance, Lee-Kim, & Ardila-Rey, 2001; Rutland, Killen, & Abrams, 2010). From this perspective, children's inferences about how people will behave (in the present studies) are in contrast to their own moral beliefs about how people should behave. Yet, despite preschoolers' general commitment to fairness, the possibility that children view people as having special moral obligations to their own group members cannot be entirely ruled out. This possibility is consistent with several theoretical accounts of morality proposed by social and cultural psychologists (Cohen, Montoya, & Insko, 2006; Dovidio, 1984; Haidt & Joseph, 2007; Haidt & Kesebir, 2010; Levine, Cassidy, Brazier, & Reicher, 2002; Levine & Thompson, 2004), and there is recent developmental data that appear consistent with this possibility (Castelli, De Amicis, & Sherman, 2007; Rhodes & Brickman, 2011). Thus, this remains an important area for future work.

As noted earlier, these three possibilities are not mutually exclusive and future work should examine the extent to which each contributes to the development of children's expectations of between-group harm. Each of the proposals above could be specific to representations of *harm*; thus, any of these accounts would be consistent with the possibility that group-based inferences about helping involve additional processes and a more protracted developmental time frame. Future work should examine how children's social-cognitive skills and experiences in groups relate to their predictions about helpful and harmful behaviors across the preschool and early school-age years, to document more precisely whether different developmental processes underlie these expectations. Future work should also examine how children's rudimentary naïve theories of social groups—as documented in

the present studies—contribute to the richer and more nuanced theories of older children and adults (Abrams et al., 2008, 2009; Bigler et al., 2008).

Understanding how children predict human action is a central component of research on conceptual development (Wellman & Gelman, 1992). Whereas the majority of research in this area has examined how children appeal to individual mental states to make these predictions, there has recently been increasing emphasis on understanding how children make these predictions by reference to social causes that extend beyond the individual, including social categories, norms, and morality (Hirschfeld, 1996; Olson & Dweck, 2008; Wellman & Miller, 2008). This emphasis—on considering children's *naïve sociology* along with their *naïve psychology*—is particularly important given that preschool-age children often weight the causal features specified by naïve sociology (e.g., categories, norms) more heavily than individual mental states (e.g., traits, desires) to predict individual action (Berndt & Heller, 1986; Biernat, 1991; Diesendruck & haLevi, 2006; Kalish, 2002; Kalish & Shiverick, 2004; Lawson & Kalish, 2006; Rhodes & Gelman, 2008; Taylor, 1996). A full understanding of the development of social cognition will require examining how children consider a wide range of causal mechanisms, and specifying how they select and weight various mechanisms across development and in different social situations. The present studies suggest that naïve theories of social groups are important contributors to children's understanding of human behavior from at least the preschool years onward.

## References

- Aboud, F. E. (1988). *Children and prejudice*. London: Blackwell.
- Aboud, F. E. (2003). The formation of in-group favoritism and out-group prejudice in young children: Are they distinct attitudes? *Developmental Psychology, 39*, 48–60. doi:10.1037/0012-1649.39.1.48
- Abrams, D., & Rutland, A. (2008). The development of subjective group dynamics. In S. Levy & M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 47–65). Oxford, England: Oxford University Press.
- 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*, 1840–1856. doi:10.1046/j.1467-8624.2003.00641.x
- Abrams, D., Rutland, A., Ferrell, J., & Pelletier, J. (2008). Children's judgments of disloyal and immoral peer behavior: Subjective group dynamics in minimal intergroup contexts. *Child Development, 79*, 444–461. doi:10.1111/j.1467-8624.2007.01135.x
- Abrams, D., Rutland, A., Ferrell, J., & Pelletier, J. (2009). Children's group nous: Understanding and applying peer exclusion within and between groups. *Child Development, 80*, 224–243. doi:10.1111/j.1467-8624.2008.01256.x
- Benenson, J., Antonellis, T., Cotton, B., Noddin, K., & Campbell, K. (2008). Sex differences in children's formation of exclusionary alliances under scarce resource conditions. *Animal Behaviour, 76*, 497–505. doi:10.1016/j.anbehav.2008.01.027
- Benenson, J., Apostoleris, N., & Parnass, J. (1997). Age and sex differences in dyadic and group interaction. *Developmental Psychology, 33*, 538–543. doi:10.1037/0012-1649.33.3.538
- Berndt, T. J., & Heller, K. A. (1986). Gender stereotypes and social inferences: A developmental study. *Journal of Personality and Social Psychology, 50*, 889–898. doi:10.1037/0022-3514.50.5.889
- Biernat, M. (1991). Gender stereotypes and the relationship between masculinity and femininity: A developmental analysis. *Journal of Personality and Social Psychology, 61*, 351–365. doi:10.1037/0022-3514.61.3.351
- Bigler, R. (1995). The role of classification skill in moderating environmental influences on children's gender stereotyping: A study of the functional use of gender in the classroom. *Child Development, 66*, 1072–1087. doi:10.2307/1131799
- Bigler, R., Arthur, A., Hughes, J., & Patterson, M. (2008). The politics of race and gender: Children's perceptions of discrimination and the U.S. presidency. *Analyses of Social Issues and Public Policy, 8*, 1–30. doi:10.1111/j.1530-2415.2008.00161.x
- Bigler, R. S., Jones, L. C., & Lobliner, D. B. (1997). Social categorization and the formation of intergroup attitudes in children. *Child Development, 68*, 530–543.
- Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory. *Current Directions in Psychological Science, 16*, 162–166. doi:10.2307/1131676
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues, 55*, 429–444. doi:10.1111/0022-4537.00126
- Castelli, L., De Amicis, L., & Sherman, S. J. (2007). The loyal member effect: On the preference for ingroup members who engage in exclusive relations with the ingroup. *Developmental Psychology, 43*, 1347–1359. doi:10.1037/0012-1649.43.6.1347
- Cohen, T. R., Montoya, R. M., & Insko, C. A. (2006). Group morality and intergroup relations: Cross-cultural and experimental evidence. *Personality and Social Psychology Bulletin, 11*, 1559–1572. doi:10.1177/0146167206291673
- Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences, 7*, 173–179. doi:10.1016/S1364-6613(03)00057-3
- Diesendruck, G., & haLevi, H. (2006). The role of language, appearance, and culture in children's social

- category-based induction. *Child Development*, 77, 539–553. doi:10.1111/j.1467-8624.2006.00889.x
- Dovidio, J. (1984). Helping behavior and altruism: An empirical and conceptual overview. *Advances in Experimental Social Psychology*, 17, 361–427.
- Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of “minimal” group affiliations in children. *Child Development*, 82, 793–811. doi:10.1111/j.1467-8624.2011.01577.x
- Gelman, S., Collman, P., & Maccoby, E. (1986). Inferring properties from categories versus inferring categories from properties: The case of gender. *Child Development*, 57, 396–404. doi:10.2307/1130595
- 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. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Foundations and the future, evolution and cognition* (pp. 367–391). New York: Oxford University Press.
- Haidt, J., & Kesebir, S. (2010). Morality. In S. T. Fiske, D. Gilbert, & G. Lindzey (Eds.), *The handbook of social psychology* (5th ed., pp. 797–832). Hoboken, NJ: Wiley.
- Hirschfeld, L. A. (1995). Do children have a theory of race? *Cognition*, 54, 209–252. doi:10.1016/0010-0277(95)91425-R
- Hirschfeld, L. A. (1996). *Race in the making*. Cambridge, MA: MIT Press.
- Kalish, C. W. (2002). Children’s predictions of consistency in people’s actions. *Cognition*, 84, 237–265. doi:10.1016/S0010-0277(02)00052-5
- Kalish, C., & Lawson, C. (2008). Development of social category representations: Early appreciation of roles and deontic relations. *Child Development*, 79, 577–593. doi:10.1111/j.1467-8624.2008.01144.x
- Kalish, C., & Shiverick, S. (2004). Rules and preferences: Children’s reasoning about motives for behavior. *Cognitive Development*, 19, 401–416. doi:10.1016/j.cogdev.2004.05.004
- Katz, P., Katz, I., & Cohen, S. (1976). White children’s attitudes towards Blacks and the physically handicapped: A developmental study. *Journal of Educational Psychology*, 68, 20–24. doi:10.1037/0022-0663.68.1.20
- Killen, M. (2007). Children’s social and moral reasoning about exclusion. *Current Directions in Psychological Science*, 16, 32–36. doi:10.1111/j.1467-8721.2007.00470.x
- Killen, M., Pisacance, K., Lee-Kim, J., & Ardila-Rey, A. (2001). Fairness or stereotypes? Young children’s priorities when evaluating group exclusion and inclusion. *Developmental Psychology*, 37, 587–596. doi:10.1037/0012-1649.37.5.587
- Kinzler, K., & Shutts, K. (2008). Memory for “mean” over “nice”: The influence of threat on children’s face memory. *Cognition*, 107, 775–783. doi:10.1016/j.cognition.2007.09.005
- Kinzler, K., Shutts, K., DeJesus, J., & Spelke, E. (2009). Accent trumps race in guiding children’s social preferences. *Social Cognition*, 27, 623–634. doi:10.1521/soco.2009.27.4.623
- Kircher, M., & Furby, L. (1971). Racial preferences in young children. *Child Development*, 42, 2076–2078. doi:10.2307/1127610
- Lawson, C. A., & Kalish, C. W. (2006). Inductive inferences across time and identity: Are category members more alike than single individuals? *Journal of Cognition and Development*, 7, 233–352. doi:10.1207/s15327647jcd0702\_5
- Levine, M., Cassidy, C., Brazier, G., & Reicher, S. (2002). Self-categorization and bystander non-intervention: Two experimental studies. *Journal of Applied Social Psychology*, 32, 1452–1463. doi:10.1111/j.1559-1816.2002.tb01446.x
- Levine, M., & Thompson, K. (2004). Identity, place, and bystander intervention: Social categories and helping after natural disasters. *Journal of Social Psychology*, 144, 229–245. doi:10.3200/SOCP.144.3.229-245
- Maccoby, E., & Jacklin, C. (1987). Gender segregation in childhood. In H. W. Reese (Ed.), *Advanced in child development and behavior* (Vol. 20, pp. 239–287). San Diego, CA: Academic.
- Nesdale, D. (2004). Social identity processes and children’s ethnic prejudice. In M. Bennett & F. Sani (Eds.), *The development of the social self* (pp. 219–246). London: Psychology Press.
- Nesdale, D., & Flessner, D. (2001). Social identity and the development of children’s group attitudes. *Child Development*, 72, 506–517. doi:10.1111/1467-8624.00293
- Olson, K., & Dweck, C. (2008). A blueprint for social cognitive development. *Perspectives on Psychological Science*, 3, 193–202. doi:10.1111/j.1745-6924.2008.00074.x
- Patterson, M. M., & Bigler, R. S. (2006). Preschool children’s attention to environmental messages about groups: Social categorization and the origins of intergroup bias. *Child Development*, 77, 847–860. doi:10.1111/j.1467-8624.2006.00906.x
- Rhodes, M. (in press). The social allegiance hypothesis. In M. Banaji & S. A. Gelman (Eds.), *The development of social cognition*. New York: Oxford University Press.
- Rhodes, M., & Brickman, D. (2011). The influence of competition on children’s social categories. *Journal of Cognition and Development*, 12, 194–221. doi:10.1080/15248372.2010.535230
- Rhodes, M., & Gelman, S. A. (2008). Categories influence predictions about individual consistency. *Child Development*, 79, 1270–1287. doi:10.1111/j.1467-8624.2008.01188.x
- Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology*, 59, 244–274. doi:10.1016/j.cogpsych.2009.05.001
- Rutland, A., Killen, M., & Abrams, D. (2010). A new social-cognitive developmental perspective on prejudice: The interplay between morality and group identity. *Perspectives on Psychological Sciences*, 5, 279–291. doi:10.1177/1745691610369468
- Sherif, M., Harvey, O., White, B., Hood, W., & Sherif, C. (1961). *Intergroup conflict and cooperation: The robbers cave*

- experiment. Norman: University of Oklahoma Book Exchange.
- Shutts, K., Roben, C., & Spelke, E. (in press). Children's use of social categories in thinking about people and social relationships. *Journal of Cognition and Development*.
- Spielman, D. A. (2000). Young children, minimal groups and dichotomous categorization. *Personality & Social Psychology Bulletin*, 26, 1433–1441. doi:10.1177/0146167200263010
- Taylor, M. G. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development*, 67, 1555–1571. doi:10.2307/1131718
- Taylor, M., Rhodes, M., & Gelman, S. (2009). Boys will be boys; cows will be cows: Children's essentialist reasoning about gender categories and animal species. *Child Development*, 80, 461–481. doi:10.1111/j.1467-8624.2009.01272.x
- Waxman, S. (2010). Names will never hurt me? Naming and the development of racial and gender categories in preschool-aged children. *European Journal of Social Psychology*, 40, 593–610. doi:10.1002/ejsp.732
- Wellman, H. (2002). Understanding the psychological world: Developing a theory of mind. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 167–187). Malden, MA: Blackwell.
- Wellman, H., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72, 655–684. doi:10.1111/1467-8624.00304
- Wellman, H., & Gelman, S. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology*, 43, 337–375. doi:10.1146/annurev.ps.43.020192.002005
- Wellman, H., & Miller, J. (2008). Including deontic reasoning as fundamental to theory of mind. *Human Development*, 51, 105–135. doi:10.1159/000115958
- Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, 69, 1–34. doi:10.1016/S0010-0277(98)00058-4

### Supporting Information

Additional supporting information may be found in the online version of this article:

**Appendix S1.** Text and Images for Introductory Stories and Sample Test Question.

**Appendix S2.** Probabilities of Between-group Predictions by Test Question Setting, Study 1.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.