Running head: CATEGORIES AND OBLIGATIONS

Social categories as markers of intrinsic interpersonal obligations

Marjorie Rhodes and Lisa Chalik

New York University

Please address correspondence to: Dr. Marjorie Rhodes, 6 Washington Place, rm. 301, New York, NY 10003; marjorie.rhodes@nyu.edu; 212-998-3546; fax: 212-995-4866.

Abstract

Social categorization is an early-developing feature of human social cognition, yet the role that social categories play in how children understand and predict human behavior has been unclear. These studies test whether a foundational functional role of social categories is to mark people as intrinsically obligated (e.g., to protect, not to harm) to one another. In three studies, children (ages 3-9, $N = 124$) viewed only within-category harm as violating intrinsic obligations; in contrast, they viewed between-category harm as violating extrinsic obligations defined by explicit rules. These data indicate that children view social categories as marking patterns of intrinsic interpersonal obligations, suggesting that a key function of social categories is to support inferences about how people will relate to members of their own and other groups.

Keywords: Social categorization, Cognitive development, Intuitive theories

Social categories as markers of intrinsic interpersonal obligations

From early infancy, children have countless experiences with human behaviors and human variation. Classifying people into categories (e.g., girls, French-speakers) is a crucial way of organizing these experiences. Children form categories based on familiar social criteria (gender, race, language) within the first year of life (Bar-Haim, Ziv, Lamy, & Hodes, 2002; Kinzler, Dopoux, & Spelke, 2007; Quinn, Yahr, Kuhn, Slater, & Pascalis, 2002) and by a wide-range of flexible criteria in early childhood (Bigler & Liben, 2007; Dunham, Baron, & Carey, 2011; Patterson & Bigler, 2006).

What role do these social categories play in how children understand and predict their environment? One perspective is that children hold an intuitive theory that social categories—much like animal categories—mark individuals who are fundamentally similar to each other (Hirschfeld, 1996). On this account, this intuitive theory facilitates the extension of information about one category member (e.g., that one girl likes a particular new game) to other members of the category (e.g., a prediction that other girls will like it too; Diesendruck & haLevi, 2006). Indeed, young children use some social categories, especially gender categories, to make these kinds of predictions about individual behavioral and psychological characteristics (Gelman, Collman, & Maccoby, 1986; Waxman, 2010). Thus, an intuitive theory that social categories mark fundamental similarities supports a powerful mechanism for using social categories to predict human behavior.

Yet, there are several reasons to suspect that this account does not capture the full functional role that social categories hold in early social cognition. First, by preschool, children can categorize people in many ways (e.g., by gender, race, language, teams, shirt colors, and so on), and are sensitive to many of these categorical distinctions in their feelings and behaviors

towards their own group members (Bigler & Liben, 2007; Dunham et al., 2011; Maccoby & Jacklin, 1987); yet, children use only a small subset of categories to make the types of inferences described above. For example, although children can categorize based on race and show race-based social preferences when tested in experimental contexts (Aboud, 1988; Baron & Banaji, 2006; Dunham, Baron, & Banaji, 2008), they do not view race as marking people who are fundamentally similar to each other (Kinzler & Dautel, 2012; Rhodes & Gelman, 2009) or use race to predict individual psychological or behavioral characteristics (Rhodes, in press, a; Shutts, Roben, & Spelke, in press). Similarly, preschool-age children can learn novel social categories based on shirt colors and labels and are sensitive to these categories in their own feelings and behaviors (Dunham et al., 2011; Patterson & Bigler, 2006), but do not use them to predict psychological properties (e.g., they do not expect members of those groups to share preferences for the same games; Kalish, 2012). Thus, many categories to which children are sensitive do not appear to invoke a naïve theory that social categories mark fundamental similarities.

Second, the view that social categories mark fundamental similarities does not provide a framework for understanding some of the uniquely *social* functions that categories of people might serve. For example, in everyday life, social categories may serve to predict patterns of social relationships and social interactions—who will be friends vs. enemies, cooperate vs. compete, or help vs. harm each other. Using social categories to make these types of inferences would rely not an assumption that category members are fundamentally similar to each other, but on beliefs about how category memberships constrain how people relate to one another. In this way, social categories could be used to predict patterns of social structure (how members of a category relate to one another and to members of other groups), but not necessarily the individual characteristics of specific members. Consistent with this possibility, children use race-based

categories to predict social relationships (e.g., who will be friends with each other) at a younger age than they use them to predict individual characteristics (who will share preferences; Shutts et al., in press). Also, although children do not expect novel categories based on clothing or labels to indicate shared psychological properties (Kalish, 2012; Kalish & Lawson, 2008), children age three and older will use such categories to predict patterns of specific social interactions. In particular, children use such categories to predict harmful interactions, expecting agents to harm (e.g., hit, tease) members of contrasting categories instead of members of the agents' own (Rhodes, in press, b).

Children's use of social categories to predict social interactions and social relationships does not follow from an intuitive theory that social categories mark individuals who are fundamentally similar to each other. Yet, the nature of the intuitive theory that *could* underlie such inferences remains unknown. Without specifying this intuitive theory, we cannot predict the types of inferences that social categories will support nor identify the implications of these inferences for social cognition more generally. Given that social categorization is an early-emerging and robust component of social cognition with multiple cognitive and behavioral consequences, identifying the nature of the intuitive theory that shapes the acquisition and use of such categories is a critical challenge for research on the development of social cognition.

The present studies test the proposal that children have a second intuitive theory of social categories—in particular, an intuitive theory that social categories mark people who hold intrinsic interpersonal obligations to one another. On this account, children have systematic, abstract expectations that people are intrinsically obligated to protect, not to harm, and to support members of their own groups, and the inferences described above (Rhodes, in press, b; Shutts et al., in press) reflect expectations that people will behave in line with these obligations (Kalish &

Shiverick, 2004). This intuitive theory would facilitate inferences about how people will relate to one another and thus provides another powerful way that social categories could be used to predict human action.

The present research directly tests whether young children view categories as marking people who are intrinsically obligated to each other. To do so, we built on methods from previous work on moral development (Smetana, 1981; Turiel, 1983), which tested beliefs about intrinsic obligations by examining judgments about whether the wrongness of obligation-violations is dependent on the presence of explicit rules. Using this method, when children maintain that an action (e.g., hitting someone) is wrong, even if there are no rules in the agent's environment prohibiting the action (e.g., no rules against hitting), this suggests that they view the action as violating an intrinsic obligation—an obligation that exists regardless of the external environment. In contrast, if children view the wrongness of an action as dependent on rules, this indicates that they view it as violating an obligation that exists only in the context of explicit social agreements. Thus, applying this method to the present context, we test whether children view people as intrinsically obligated not to harm their own category members, but as prohibited from harming members of other categories only in the presence of explicit social rules.

Study 1

Participants

Participants included 23 preschool-age children (*M* age = 4.5 years, range = 4.0-4.9; 14 male, 9 female; 39% White, 4% African American, 22% Asian American, 22% Hispanic, 4% multi-ethnic, remainder unknown) recruited from and tested at the Children's Museum of Manhattan. An additional eight children were tested but excluded (four for disruptions during the testing session and four for failing to meet inclusion criteria, see below).

Procedure

Participants were introduced to two novel categories through a brief story. Novel categories were used to test whether children have abstract beliefs about how categories mark obligations, which are observable as children make sense of new social divisions and are not dependent on knowledge of specific group customs or histories. Using novel groups also allows us to examine responses to groups for which children themselves do not hold membership, thus examining children's abstract beliefs, without involving generalized biases in favor of their own group members.

The story introduced two categories—the "Flurps" and the "Zazes"—that were marked by shirt color (e.g., red team and blue team). To ensure that children treated the novel categories as meaningful, the groups were described as engaging in within-group cooperation. An activity was briefly described in which the members of each group worked together to build a block tower (scripts and images are available in Rhodes, in press, b, Study 1). The activities were not competitive and no interactions between members of the different groups were described.

Next, children were told about a harmful interaction that took place between two individuals on a playground. Children heard a scenario in which a perpetrator harmed a member of his or her own category (within-group harm, "One day, a Zaz teased another Zaz and hurt his feelings"), and a scenario in which a perpetrator harmed a member of the other category (between-group harm, "One day, a Zaz teased a Flurp and hurt his feelings") in counter-balanced order. One item involved teasing, and the other involved social exclusion, with assignment of the scenario (teasing, exclusion) to Harm Context (within-group harm, between-group harm) counter-balanced across participants. There were no effects of scenario-type, so data were

collapsed across this distinction to focus on the comparison of within-group and between-group harm.

To measure whether children viewed the harmful action as violating intrinsic obligations, children were asked a series of questions. Immediately after the event was described, children were asked**:** "Was what the Zaz did OK or not OK? [if not OK] Was it a little bad, pretty bad, or very, very bad? (Scored 0 = OK, 1 = A little bad, 2= Pretty bad, 3 = Very, very bad)." Then, we provided information about the characters' environment, "What if there was no rule in their school against teasing? Let's pretend that in the school they go to the teachers said that the kids could tease each other. Then would it be OK or not OK for the Zaz to tease [another Zaz/ a Flurp] and hurt his feelings? [If not OK] Would it be a little bad, pretty bad, or very, very bad?". If children view people as intrinsically obligated only to their own category members, they should maintain that it is wrong for a perpetrator to harm a member of his or her own group regardless of explicit rules (thus, their ratings should not change across questions), but should view the wrongness of harming a member of the other group as contingent on explicit rules. Because this study was designed to examine beliefs about the wrongness of harmful actions—whether wrongness stems from intrinsic or extrinsic causes—children were included only if they initially identified the harmful action as unacceptable (four children were excluded for initially identifying the action as "okay"). Following previous developmental work, children were asked to explain their evaluations (see Table S1 in the online supplementary materials).

Results and Discussion

Children's evaluations were analyzed through a 2 (Harm Context: Within-group, Between-group) X 2 (Rule Information: Before, After) repeated measures analysis of variance, with both factors as within-subjects variables. Children gave harsher ratings before (M = 2.30,

SE = .17) than after (*M* = 1.70, *SE* = .22) the rule information, $F(1, 22) = 5.69$, $p = .03$, partial $\eta^2 = .20$. As predicted, however, they did so only for between-group harm; ratings of between-group harm became less negative following the rule information, $p = .001$, Cohen's $d = .92$, whereas ratings of within-group harm did not change, $p > .50$, see Figure 1. The Harm Context X Rule Information interaction was reliable, $F(1, 22) = 7.06$, $p = .01$, partial $\eta^2 = .24$. Thus, four-year-olds treated only within-group harm as violating intrinsic obligations.

Study 2

The purpose of Study 2 was rule out the possibility that children only consider categories as marking intrinsic obligations in contexts that might be conducive to between-group competition. Although no competition was described in Study 1, the groups were referred to as "teams", which could have led children to infer the presence of a contest.

Participants

Participants were 63 children recruited from and tested at private preschools (38 male, 25 female; *M* age = 4.63 years, range = 3.55 – 5.98; 35% multi-ethnic, 40% Asian American; remainder unknown). Because the age range of Study 2 was larger than the range of Study 1, spanning the entire preschool period, we divided children into two age-groups for analyses to test whether a similar pattern held across this period: Younger preschoolers (*N* = 33, *M* = 4.14, range = 3.55-4.50) and Older preschoolers (*N* = 30, *M* = 5.15 years, range = 4.60-5.98).

Procedures

To provide a thorough test of whether children have a robust expectation that groups mark intrinsic obligations even in the absence of between-group competition, Study 2 included two conditions. The first condition, referred to as "Cooperation," replicated Study 1, but with the groups referred to simply as "groups" instead of as "teams."

The second condition, referred to as "Language", examined whether the pattern found in Study 1 holds across another group context, one in which children are likely to view the groups as meaningful based on their prior knowledge (Kinzler, Shutts, DeJesus, & Spelke, 2009), but where no within-group cooperation was described at all. Children were shown four characters from each group, and told, "Here are the Flurps! Here are the Zazes! I want to show you what they sound like. Let's listen to some things they say." The experimenter pointed individually to two members of each group and for each played an audio recording of a child speaking (Kinzler et al., 2009). The members of one of the groups spoke French, whereas the members of the other group spoke English. The content of the audio recording was neutral and identical across languages (e.g., "hide and seek is a fun game to play").

After introduction to the novel groups, children were asked the test questions used in Study 1 about four separate scenarios (2 within-group harm and 2 between-group harm; one harmful action in each context involved teasing, as in Study 1, the other involved denying someone access to resources; e.g., "The Flurp took all of the crackers for himself and wouldn't share them with the Zaz. The Zaz didn't have any crackers and was sad"), with order of Harm Context counter-balanced across participants. There were no effects of scenario type, so we collapsed across this distinction to focus on comparisons of within-group and between-group harm.

Results and Discussion

We conducted a 2 (Age-group: Older preschoolers, Younger preschoolers) X 2 (Harm Context: Between-group, Within-group) X 2 (Rule Information: Before, After) X 2 (Condition: Cooperation, Language) repeated measures analysis of variance, with Harm Context and Rule Information as within-subjects factors. Children rated the actions more negatively before (M =

2.25, *SE* = .07) than after (*M* = 1.87, *SE* = .13) the rule information, $F(1, 59) = 10.23$, $p = .002$, partial $\eta^2 = .15$. As in Study 1, however, they did so only for between-group harm (Before, *M* = 2.23, *SE* = .09, After, *M* = 1.77, *SE* = .14), $p < .001$, Cohen's $d = .56$. Ratings for within-group harm did not change (Before, *M* = 2.20, *SE* = .09, After, *M* = 1.97, *SE* = .14, $p > .30$). The Harm Context X Rule Information interaction was reliable, $F(1, 59) = 4.22$, $p = .04$, partial $\eta^2 = .07$, and there were no main or interactive effects of Age-group or Condition. Inspection of the means confirmed that identical patterns were found across both age-groups and conditions. Thus, Study 2 confirms that preschool-age children have a robust expectation that only within-group harm violates intrinsic obligations, which they apply even in the absence of between-group competition and across multiple ways of defining social categories.

Study 3

The aim of Study 3 was to examine whether the obtained pattern held across older childhood. Kalish and Lawson (2008; also Kalish, 2012) suggest that obligations are particularly central to younger children's understanding of social categories. Thus, one possibility is that the effects documented in Studies 1-2 are specific to the early childhood period. Yet, another possibility is that the intuitive theory that social categories mark intrinsic obligations continues to shape children's understanding of social categories across childhood development. Thus, here we test whether older children view categories as marking patterns of intrinsic interpersonal obligations.

Participants and Procedures

Participants included 38 children ages 7-9 (19 male, 19 female; *M* age = 8.5 years, range = 7.2 - 9.8 years; 67% White, 3% Asian, 7% Hispanic, remainder unknown), recruited from and tested at a public elementary school in New York City (two additional children were tested but

excluded for failing to meet the inclusion criteria described in Study 1). Due to time constraints on the research sessions, children completed only one set of items each. Participants were introduced to the novel categories as in Study 1, and then were randomly assigned to a scenario about either within-group harm or between-group harm (the "teasing item" used in Studies 1 and 2).

Results and Discussion

Children rated the action as worse before (M = 2.10, SE = .14) than after (M = 1.37, SE = .21) the rule information, $F(1, 36) = 11.40$, $p = .002$, partial $\eta^2 = .24$, and within-group harm (M = 2.03, SE = .19) as worse than between-group harm (M = 1.44, SE = .20), $F(1, 36) = 4.48$, p = 04, partial $\eta^2 = .11$. The interaction between rule information and group context was reliable, $F(1, 36) = 4.48$, $p = .03$, partial $\eta^2 = .12$. As shown in Figure 2, the pattern was identical as was found in previous studies, ratings of within-group harm did not change across questions, $p > .40$, but ratings of between-group harm became less negative following the rule information, $p <$ .001, Cohen's D = 1.20. Thus, older children, like younger children in Studies 1-2, treat social categories as marking people who are intrinsically obligated to one another.

General Discussion

We found that children (ages 3-9) view social categories as marking patterns of intrinsic interpersonal obligations; that is, that they view people as intrinsically obligated only to their own group members. In three studies, children viewed within-group harm as wrong regardless of explicit rules, but viewed the wrongness of between-group harm as contingent on the presence of such rules.

We suggest that the intuition that categories mark patterns of intrinsic obligations underlies an important functional role of social categories in social cognition, by supporting

predictions of obligation-relevant behaviors. In this manner, children can use social categories to predict elements of social structure—how people will interact with each other, the relationships they will form, and to whom they will direct harmful behaviors. In the present studies, children demonstrated these intuitions both for completely novel and arbitrary social categories for which they themselves did not hold membership, as well as for categories based on more familiar distinctions (language differences). Thus, these data suggest that abstract expectations that social categories mark patterns of intrinsic interpersonal obligations are readily elicited and robustly relied on across different contexts.

The present studies primarily examined novel groups, marked by labels and shirt colors, in order to test children's abstract beliefs about how categories shape obligations independent of their knowledge of previous group histories or their own category memberships. A critical question for future work is how these effects extend to familiar categories that children might encounter in their everyday lives, including those about which they have more background knowledge and for larger social groups in which the individual members may not be personally familiar with one another. A particularly important issue to address is whether the present findings hold for categories that are not marked by labels. Labeling highlights categories and facilitates category-based reasoning for both social (Bigler & Liben, 2007; Gelman & Heyman, 1999; Waxman, 2010) and non-social (for review, see Waxman, 1999) categories and plays an especially important role as children learn new categories. Thus, the present findings may indeed have depended on the labels provided for these novel categories.

Nevertheless, the present findings go far beyond the general effects of labeling. Although labels facilitate the development of categorization across domains, the types of inferences that labeled categories support vary by domain. For animal categories, for example, children treat

labels as marking individuals who are similar to each other in non-obvious ways (Gelman, 2003). For novel social categories, however, labels do not support inferences that individual members are similar to each other (e.g., that category members will share non-obvious preferences or behaviors; Kalish, 2012), even after extensive repeated exposure to such labels (Rhodes, Leslie, & Tworek, 2012). Instead, children use novel labeled social categories to evaluate how people should relate to one another, as shown in the present studies, following a fairly brief introduction to the categories. Thus, the phenomenon revealed here may result from the interplay between the general effects of labeling with children's intuitive theory of how categories shape the social world.

This study examined responses to hypothetical, explicit alterations of familiar moral rules. As shown by previous work on moral development (Smetana, 1981; 2006), by age three, children generally view these actions that cause harm (e.g., teasing) as wrong and prohibited. Use of hypothetical situations was necessary to prevent children from developing the belief that these harmful behaviors are actually permissible in certain contexts. Yet, this approach raises important questions regarding whether the present findings generalize to more realistic events. Critically, the instructions that signaled the hypothetical nature of the rule change (e.g., "let's pretend…" and "what if") were held constant across scenarios describing within-group and between-group harm and thus cannot account for why children showed more lenient evaluations of harm only on a subset of items (i.e., only for between-group harm). Also, children often reason quite similarly about hypothetical harmful transgressions and transgressions that they actually view in their environment (Turiel, 2008). Still, the generalizability of findings using hypothetical terms is an important issue to address in future work

The key aim of this research was to examine the role of obligations in representations of social categories. Yet, because we examined how children evaluate harmful actions, the present data may also have implications for the development of moral judgment. In previous work on moral development, viewing a transgression as wrong regardless of external rules has been taken as a key hallmarks of moral judgment, whereas viewing the wrongness of a transgression as contingent on rules has been taken to indicate non-moral, conventional reasoning (Smetana, 1981; Turiel, 1983). Applying this distinction to the present context would thus indicate that children view within-group harm as a moral transgression, but view between-group harm as breaking only conventionalized rules. In this manner, social groups may operate to define "moral boundaries," consistent with some anthropological, social psychological, cognitive neuroscience, and philosophical theories (Cohen, Montoya, & Insko, 2006; Greene, 2003; Haidt, 2008; Meier & Hinsz, 2004; Rai & Fiske, 2011; Shweder, Mahapatra, & Miller, 1990).

The present studies show that children treat social categories as marking not only whether an individual is bound to specific conventional norms (e.g., regarding foods or dress; Kalish, 2012; Kalish & Lawson, 2008) but also to one of the most fundamental moral obligations—the obligation not to harm. These findings, together with prior work, indicate that children hold a powerful intuitive theory that social categories mark how people ought to behave. We propose that this intuition forms the bases of naïve sociology—children's abstract expectations about the structure of the social world—and thus contributes to and guides the acquisition of social knowledge.

Acknowledgements

Funding was provided by NSF grant BCS-1226942 to Rhodes. We thank the New York City Department of Education, the Children's Museum of Manhattan, and the preschools that participated in this research; Noey Neumark, Kelli Grobe, Dhaya Ramarajan, and Kimberly Kowalski for assistance with data collection; Dr. Katherine Kinzler for the audio files for Study 2; Drs. Gregory Murphy, Kristina Olson, Joshua Knobe for helpful feedback on a previous draft; and Karl Edwards for study illustrations.

References

Aboud, F. E. (1988). *Children and prejudice*. London: Blackwell.

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. (2006). Nature and Nurture in Own-Race Face Processing. *Psychological Science*, *17*, 159-163.

Baron, A.S., & Banaji, M.R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages, 6, 10, and adulthood. *Psychological Science, 17,* 53-58.

Bigler, R.S., & Liben, L.S. (2007). Developmental intergroup theory. *Current Directions in Psychological Science, 16,* 162-166.

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.

Diesendruck, G., & haLevi, H. (2006). The role of language, appearance, and culture in children's social category-based induction. *Child Development, 77,* 539-553.

Dunham, Y., Baron, A.S., & Banaji, M.R. (2008). The development of implicit intergroup cognition. *Trends in Cognitive Sciences, 12,* 248-253.

Dunham, Y., Baron, A.S., & Carey, S. (2011). Consequences of 'minimal' group affiliations in children. *Child Development, 82,* 793-811.

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.

Greene, J. (2003). From neural 'is' to moral 'ought': what are the moral implications of neuroscientific moral psychology? *Nature Reviews: Neuroscience, 4,* 847-852.

Haidt, J. (2008). Morality. *Perspectives on Psychological Science, 3,* 65-72.

Gelman, S.A., & Heyman, G.D. (1999). Carrot-eaters and creature-believers: The effects of lexicalization on children's inferences about social categories. *Psychological Science, 10,* 489-493.

Hirschfeld, L.A. (1996). *Race in the making.* Cambridge: MIT Press.

Kalish, C. (2012). Generalizing norms and preferences within social categories and individuals. *Developmental Psychology, 48,* 1133-1143.

Kalish, C., & Lawson, C. (2008). Development of social category representations: Early appreciation of roles and deontic relations. *Child Development, 79,* 577-593.

Kalish, C. & Shiverick, S. (2004). Rules and preferences: Children's reasoning about motives for behavior. *Cognitive Development, 19,* 401-416.

Kinzler, K.D., & Dautel, J. (2012). Children's essentialist reasoning about language and race. *Developmental Science, 15,* 131-138.

Kinzler, K.D., Dupoux, E., & Spelke, E.S. (2007). The native language of social cognition. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 104,* 12577-12580.

Kinzler, K., Shutts, K., Dejesus, J., & Spelke, E. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition, 27,* 623-634.

Maccoby, E., & Jacklin, C. (1987). Gender segregation in childhood. In H.W. Reese (Ed.), *Advances in child development and behavior* (Vol. 20, pp. 239-287). San Diego, CA: Academic Press.

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.

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.

Quinn, P., Yahr, J., Kuhn, A., Slater, A., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception*, *31*, 1109-1121.

Rai, T.S., & Fiske, A.P. (2011). Moral psychology in relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review, 118,* 57-75.

Rhodes, M. (in press, a). The social allegiance hypothesis. Chapter to appear in M. Banaji & S.A. Gelman (Eds.) *The Development of Social Cognition*.

Rhodes, M. (in press, b). Naïve theories of social groups. *Child Development.*

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.

Rhodes, M., Leslie, S.J., & Tworek, C. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences.* Published online before print: August 6, 2012.

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.*

Shweder, R., Mahapatra, M., & Miller, J. (1990). Culture and moral development. In J. Stigler, R. Shweder, & G. Herdt (Eds.), *Cultural psychology: Essays on comparative human development* (pp. 130-204). New York: Cambridge University Press.

Smetana, J. (1981). Preschool children's conceptions of moral and social rules. *Child Development, 52,* 1333-1336.

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* (pp. 119-154). Mahwah, NJ: Erlbaum.

Turiel, E. (1983). *The development of social knowledge: Morality and convention.* Cambridge: Cambridge University Press.

Turiel, E. (2008). Thoughts about actions in social domains: Morality, social conventions, and social interactions. *Cognitive Development, 23,* 136-154.

Waxman, S.R. (1999). The dubbing ceremony revisited: Object naming and categorization in infancy and early childhood. In D.L. Medin & S. Atran (Eds.), *Folkbiology* (pp. 233-284). Cambridge: The MIT Press.

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.

Figure Captions.

Figure 1. Four-year-olds' evaluations of harm (0 = Okay, 3 = Very, very bad) by Harm Context and Rule Information, Study 1; Error bars represent one standard error of the mean.

Figure 2. Older children's evaluations of harm (0 = Okay, 3 = Very, very bad) by Harm Context and Rule Information, Study 3; Error bars represent one standard error of the mean.