How Two Intuitive Theories Shape the Development of Social Categorization

Marjorie Rhodes

New York University

ABSTRACT—*Social categorization provides a valuable mechanism for explaining and predicting human behavior, yet also contributes to the development of social stereotyping and prejudice. Thus, understanding how social categorization develops is critical for both cognitive and social development. This article presents a theoretical perspective on the development of social categorization— that children map intuitive theories about the structure of the social world onto categories they encounter in their environment. In particular, 2 intuitive theories—that social categories are natural kinds and that social categories mark people who are obligated to one another—are described as shaping the acquisition of social categories. The article discusses implications for how children explain, predict, and interact with their environment.*

KEYWORDS—*social categorization; folk theories; social cognition*

From early infancy, children have countless experiences with human behaviors and human variation. Classifying people into categories (e.g., girls, doctors, babies) is a crucial way of organizing these experiences. Social categorization enables children to encode and retrieve information about people efficiently and provides a valuable mechanism for predicting and explaining human action (for review, see Kinzler, Shutts, & Correll, 2010). Yet, social categorization also contributes to a range of negative psychological and social phenomena, including social stereotyping and prejudiced attitudes (Bigler & Liben, 2007; Rutland,

Marjorie Rhodes, Department of Psychology, New York University.

Correspondence concerning this article should be addressed to Marjorie Rhodes, Department of Psychology, New York University, 6 Washington Place, New York, NY 10003; e-mail: marjorie. rhodes@nyu.edu.

Killen, & Abrams, 2010). Thus, social categorization is critical to our understanding of both cognitive and social development, and provides insight into important connections between these fields. This article outlines a theoretical perspective on this intriguing and socially important area.

Categories reflect domain-specific intuitive theories about the structure of the world (Murphy & Medin, 1985). For example, given the many ways that animals could be categorized (fur color, size, location, behavior), the decision to classify into species kinds (e.g., dogs vs. cats) reflects abstract beliefs about the structure of the biological world—that the biological world is composed of discrete kinds that are determined by biological inheritance (Atran, 1998). Thus, to describe the development of social categorization, it is necessary to identify the nature of children's theories about the structure of the social world. This article describes two intuitive theories of the social world—that social categories are natural kinds and that social categories mark patterns of social obligations—and examines the role of each in shaping the development of social categorization across childhood.

INTUITIVE THEORY 1: SOCIAL CATEGORIES AS NATURAL KINDS

Hirschfeld (1996) proposed that an intuitive theory that social categories are *natural kinds*—that is, that the social world is composed of discrete, coherent kinds determined by nature— guides the development of social categorization (also, Rothbart & Taylor, 1992). From this perspective, children view social category memberships as determined by birth, stable, and predictive of a wide range of physical and behavioral properties, much like animal species. By preschool (ages 3–5), children's beliefs about at least one social category—gender—indeed reflect theoretical commitments indicative of treating social categories as natural kinds. Preschoolers view gender as: (a) marking objective structure (e.g., judging that it is *wrong* to consider a boy and a girl the same kind of person, even if they have other traits in common; Rhodes & Gelman, 2009a; Rhodes, Gelman, &

Karuza, in press); (b) established by birth, stable, and conferring an innate potential (e.g., responding that being born a girl inevitably means growing up to prefer ballet to football, regardless of the environment; Taylor, Rhodes, & Gelman, 2009); and (c) identifying people who are fundamentally similar (e.g., inferring that people of the same gender will share physical, behavioral, and psychological properties, even if they are perceptually dissimilar or have different personalities; Diesendruck & Halevi, 2006; Gelman, Collman, & Maccoby, 1986).

How might the intuitive theory that social categories are natural kinds guide the development of social categorization? For animal categories, children readily map their intuitive theory onto categories in the world. For example, when children encounter a new animal category, they assume that it is a distinct kind with strict boundaries (Rhodes & Gelman, 2009b; Rhodes et al., in press) and that its members are highly similar to each other (Brandone & Gelman, 2009). For a comparable process to operate for social categorization, when children encounter a social category, they should readily make a similar series of assumptions.

Yet the developmental data on social categorization do not suggest that children map their intuitive theory that social categories are natural kinds to all categories of people they encounter. Rather, young children apply natural kind beliefs to social categories quite selectively. Also, natural kind beliefs about social categories often show protracted developmental trajectories and substantial cultural variability. Children's beliefs about racial categories illustrate these points. By preschool, young children are aware of racial categories (Dunham, Baron, & Banaji, 2008), yet they appear not to treat race as marking a natural kind. Whether they do so has been the matter of some debate. Hirschfeld (1995) demonstrated that 4-year-olds understand the physical features associated with race (e.g., skin color) as inherited and stable, and from these data, proposed that young children view race as a natural kind. Yet these studies did not test whether the children treated those physical features as inductively informative or as marking fundamentally distinct kinds of people. Critically, children treat many physical features (e.g., hair color; Rhodes, in press a) as inherited and stable, but do *not* view them as marking natural kinds.

Indeed, subsequent studies have found that preschoolers do not view racial categories in this manner. For example, Rhodes and Gelman (2009a) found that although 5-year-olds can classify people by skin color, they viewed race as a subjective, flexible way of categorizing people (in contrast to gender and animal species, which they treated as objective). Similarly, Shutts, Roben, and Spelke (in press) found that preschoolers did not expect people of the same race to share novel psychological properties (see also Waxman, 2010), suggesting they do not view race as marking fundamental similarities. Thus, despite their awareness of race, preschool-age children do not map their theory that some social categories are natural kinds to race. Rather than emerging in early childhood, natural kind beliefs

about racial categories develop slowly (between the ages of 7 and 10), in a manner dependent on children's cultural context (Kinzler & Dautel, 2012; Rhodes & Gelman, 2009a). Similarly lengthy and culture-specific developmental trajectories of natural kind beliefs have been found for other social categories as well (e.g., religion, ethnicity; Birnbaum, Deeb, Segall, Ben-Eliyahu, & Diesendruck, 2010; Deeb, Segall, Birnbaum, Ben-Eliyahu, & Diesendruck, 2011).

Studies of novel categories further illustrate the distinction between recognizing particular criteria for categorizing people and treating those categories as natural kinds. Young children quickly learn novel criteria for classifying people (e.g., shirt colors or novel labels) and show sensitivity to such categories in their social attitudes (Bigler & Liben, 2007; Dunham, Baron, & Carey, 2011). Yet, children do not treat these novel categories as natural kinds. For example, Kalish and Lawson (2008) found that 4-year-olds do not expect people who share membership in novel social categories to share preferences, suggesting that they do not view such categories as marking people who are fundamentally similar to each other. Similarly, Rhodes and Brickman (2011) found that 5-year-olds do not expect category memberships marked by shirt colors and labels to be stable, inductively informative, or determined by birth, unless they receive extensive additional input about those categories (also Rhodes, Leslie, & Tworek, 2012; see below). Thus, in examining the development of social categorization, it is critical to distinguish the evidence that children are aware of categories from evidence that they treat those categories as marking natural kinds. Evidence that children treat social categories as natural kinds requires that they view category memberships as fundamental to identity—that is, as a stable, objectively accurate way of classifying people that marks patterns of fundamental similarities and differences.

What determines whether children treat a particular social category as a natural kind? This question involves two issues: (a) Why do young children hold essentialist beliefs about gender, but not other salient social categories such as race? And (b) what determines how children map these beliefs to other categories across development? Several hypotheses have been proposed for why young children hold essentialist beliefs about gender, but not other categories. First, because of the significance of gender to the processes involved in shaping evolution, concepts of gender may be particularly constrained by intuitive cognitive biases (Cosmides, Tooby, & Kurzban, 2003; Kinzler et al., 2010). In contrast, because categories based on race (or other criteria) did not play a role in shaping human evolution, they would not be constrained in this manner and would instead depend more heavily on cultural experience. Empirical support for this possibility comes from studies showing that natural kind concepts of gender develop in early childhood even in communities where older children and adults have more flexible gender beliefs, suggesting that children's own intuitive biases shape their early concepts (Rhodes & Gelman, 2009a; see also Kurzban, Tooby, & Cosmides, 2001).

Another plausible proposal, however, is that cultural input plays a crucial role in shaping how children apply natural kind beliefs to particular categories in both early childhood and across development. From this perspective, children might be exposed to the relevant cultural input for gender categories earlier than for other categories, such as race. Language may be a key form of this cultural input. In particular, hearing generic language (language that describes abstract kinds, e.g., "boys play baseball") leads 4-year-olds, as well as adults, to develop natural kind beliefs about novel social categories that they would not otherwise view in this manner (Rhodes et al., 2012). Furthermore, parents selectively produce generic language when talking to their children about social categories for which they themselves hold natural kind beliefs (Rhodes et al., 2012); thus, generic language could serve as a mechanism that facilitates the cultural transmission of natural kind beliefs about particular social categories (see also Gelman, Taylor, & Nguyen, 2004). Other cultural factors may also play a role; for example, learning that two novel categories are engaged in intergroup conflict increases beliefs that the categories reflect fundamentally distinct kinds of people (Rhodes & Brickman, 2011); experiences with social diversity appear to contribute as well (Deeb et al., 2011; Kinzler & Dautel, 2012).

In sum, children have an intuitive theory that certain social categories mark natural kinds. Children apply this theory to social categories selectively and in a manner that is partially dependent on cultural input. Although children hold natural kind beliefs about a limited number of social categories, the social categories that they *do* view in this manner are particularly important because they are often implicated in the development of social stereotyping and prejudice (Dweck, 2009; Haslam, Rothschild, & Ernst, 2002; Keller, 2005; Leslie, in press). Thus, identifying the processes that lead children to apply natural kind beliefs to social categories is critical for determining how to prevent the development of these negative social phenomena.

INTUITIVE THEORY 2: SOCIAL CATEGORIES AS MARKERS OF SOCIAL OBLIGATIONS

As reviewed above, young children are aware of many social categories that they do not view as natural kinds (ranging from those based on race to those based on novel labels and shirt colors). Yet, although children do not view these categories as natural kinds, they may still serve important, functional roles in early social cognition. Indeed, many of the social categories that adults treat as meaningful and informative—team memberships, political parties, interest groups, and so on—are not categories that they generally think of as natural kinds (Haslam, Rothschild, & Ernst, 2000). Thus, recent work has aimed to discover whether a second intuitive theory of the social world might also contribute to the development of social categorization.

To determine what such a theory might be, it is useful to consider that although the inferential role served by natural kind categories—supporting inferences that category members are similar to each other—is important, some of the ways in which social categories contribute to human behavior do not stem from within-category similarity. For example, in daily life, social categories contribute to social relationships and interactions, shaping who will be friends or enemies, cooperate or compete, or help or harm each other. Thus, another key conceptual role that social categories could serve is to support inferences about how people relate to one another.

Recent work has tested whether young children have a second intuitive theory about the structure of the social world—that social categories mark people who are obligated to one another—that shapes the development of social categorization. This intuitive theory includes abstract expectations that members of a category have intrinsic obligations to each other (e.g., to protect and not harm each other) that do not extend beyond category boundaries (see also Kalish & Lawson, 2008). This theory thus allows predictions and explanations of obligation-relevant behaviors and relationships, for example, that members of the same category will be friends (and not enemies) with each other, and that people will refrain from harming members of their own (but not necessarily of other) categories.

In support of this proposal, Rhodes (2012) found that children use social categories to predict patterns of social interactions. Children aged 3 and older were introduced to novel social categories that were marked by labels ("Flurps" and "Zazes") and t-shirt colors. At all ages, they reliably used the categories to predict social interactions. Children predicted that agents would refrain from harming (e.g., hitting, teasing) members of the agents' own category, and instead would direct harmful actions toward members of contrasting categories. Children aged 3–5 did not use the categories to predict nice behaviors; however, they predicted that agents would engage in nice behaviors (e.g., sharing, hugging) toward members of their own and other categories equally often. (Reliable predictions of within-category nice behaviors developed around age 6.)

This pattern is consistent with the proposal that children treat some social categories as marking people who are obligated to one another. Harming violates social obligations; thus, children predict that agents will refrain from harming people with whom they share category membership. Nice behaviors—although positively valued—are not obligated (Knobe, 2010); thus, a basic intuitive theory that categories mark people who are obligated to one another does not support predictions about these behaviors. In further support of this interpretation, Rhodes and Chalik (in press) found that 4-year-olds evaluate harm among members of the same category as wrong, regardless of whether there are rules in place prohibiting the harmful actions, suggesting that they view members of the same category as intrinsically obligated not to harm one another. In contrast, children's evaluations of harm between members of different categories depended on the presence of explicit rules, suggesting that they view intrinsic obligations not to harm as stopping at category boundaries.

Providing further evidence of this second intuitive theory, Shutts et al. (in press) found that preschool-age children use racial categories to predict social relationships, but not individual preferences. Also, Kalish and Lawson (2008) found that children expect novel social categories to be characterized by distinct obligations, but not distinct psychological properties.

In contrast with the substantial input required for children to treat social categories as natural kinds (reviewed above), very limited input is needed to trigger children's use of categories to predict social interactions. In one condition of Rhodes (2012), 3-year-olds did so robustly (more than 70% of the time) based only on labels and shirt colors. Thus, children readily map their intuitive theory that categories mark people who are obligated to one another onto new categories they encounter, without the need for additional linguistic input or cultural experiences. In addition to these novel categories, children also use familiar social categories in this manner, including those based on language differences (Rhodes & Chalik, in press) and race (Rhodes, in press; see also Shutts et al., in press). Thus, children appear to apply their theory that social categories mark obligations much more broadly than their theory that certain social categories are natural kinds.

RELATIONS AMONG THEORIES

How do children's two intuitive theories of the social world work together to shape social cognition across development? The first key point to consider is that each theory supports a different conceptual role of social categories. The theory that social categories mark natural kinds supports inferences that individual category members are fundamentally similar to each other (e.g., that if one girl likes a new game, other girls will like it, too; Gelman et al., 1986). In contrast, the intuitive theory that social categories mark intrinsic obligations does not support these types of inferences about the characteristics of individual category members. Instead, this theory supports inferences about how people relate to and behave toward one another (e.g., who will be friends or enemies, who will help or harm each other; Rhodes, 2012; see also, Kalish & Lawson, 2008; Shutts et al., in press). Thus, categories reflecting each of children's intuitive theories systematically support different inferences.

A second key point is that different levels of input are required to evoke each of these theories; when children learn about a category (e.g., either a novel category in the lab or a familiar category in their environment), extensive input is required to elicit the belief that the category is a natural kind (Rhodes & Gelman, 2009a; Rhodes, Leslie, et al., 2012), whereas children will infer much more readily that new categories mark patterns of social obligations. Thus, one possibility is that when children encounter a new category (again, in the lab or their everyday lives), they first assume that it marks patterns of social obligations. Over time, if children receive certain cultural input, they may also begin to develop natural kind beliefs.

If children's experiences with a category support both of these theories, they may then hold both theories about a given category simultaneously (relying on each to make different sorts of inferences; this may happen across development for categories such as race). Alternately, children might continue to represent a category only as marking patterns of social obligations (e.g., for categories such as sports fans). Whether children ever fail to hold the belief that a category marks patterns of social obligations, instead viewing it only as a natural kind, is an open question (e.g., this description could apply to gender).

SUMMARY

Children's social categories provide a window into the abstract theories they use to make sense of a highly complex social world. Children rely on at least two intuitive theories to make sense of this environment. Understanding how children map each of these theories onto categories across development and use them to understand and predict human behavior are critical to our understanding not only of conceptual development but also of the development of a range of critical social phenomena.

REFERENCES

Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences*, *31*, 547–609.

Bigler, R., & Liben, L. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science*, *16*, 162–166.

Birnbaum, D., Deeb, I., Segall, G., Ben-Eliyahu, A., & Diesendruck, G. (2010). The development of social essentialism: The case of Israeli children's inferences about Jews and Arabs. *Child Development*, *81*, 757–777.

Brandone, A. C., & Gelman, S. A. (2009). Differences in preschoolers' and adults' use of generics about novel animals and artifacts: A window onto a conceptual divide. *Cognition*, *110*, 1–22.

Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences*, *7*, 173–179.

Deeb, I., Segall, G., Birnbaum, D., Ben-Eliyahu, A., & Diesendruck, G. (2011). Seeing isn't believing: The effect of intergroup exposure on children's essentialist beliefs about ethnic categories. *Journal of Personality and Social Psychology*, *101*, 1139–1156.

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.

Dweck, C. (2009). Prejudice: How it develops and how it can be undone. *Human Development*, *52*, 371–376.

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.

Gelman, S., Taylor, M., & Nguyen, S. (2004). Mother-child conversations about gender: Understanding the acquisition of essentialist beliefs. *Monograph of the Society for Research in Child Development*, *69*, 1–145.

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology*, *39*, 113–127.

Haslam, N., Rothschild, L., & Ernst, D. (2002). Are essentialist beliefs associated with prejudice? *British Journal of Social Psychology*, *4*, 87–100.

Hirschfeld, L. A. (1995). Do children have a theory of race? *Cognition*, *54*, 209–252.

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

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

Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology*, *88*, 686–702.

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

Kinzler, K. D., Shutts, K., & Correll, J. (2010). Priorities in social categories. *European Journal of Social Psychology*, *40*, 581–592.

Knobe, J. (2010). Person as scientist, person as moralist. *Behavioral and Brain Sciences*, *33*, 315–329.

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences*, *98*, 15387–15392.

Leslie, S. (in press). The original sin of cognition: Fear, prejudice, and generalization. *Journal of Philosophy*.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289–316.

Rhodes, M. (2012). Naïve theories of social groups. *Child Development*, Advance online publication. doi:10.1111/j.1467-8624.2012.01835.x

Rhodes, M. (in press). The social allegiance hypothesis. In M. Banaji & S. A. Gelman (Eds.), *Navigating the social world: What infants, children, and other species can teach us*. 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.

Rhodes, M., & Chalik, L. (in press). Social categories as markers of intrinsic interpersonal obligations. *Psychological Science*.

Rhodes, M., & Gelman, S. A. (2009a). 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., & Gelman, S. A. (2009b). Five-year-olds' beliefs about the discreteness of category boundaries for animals and artifacts. *Psychonomic Bulletin & Review*, *16*, 920–924.

Rhodes, M., Gelman, S. A., & Karuza, J. C. (in press). Preschool ontology: The role of beliefs about category boundaries in early categorization. *Journal of Cognition and Development*.

Rhodes, M., Leslie, S. J., & Tworek, C. (2012). Cultural transmission of social essentialism. *Proceedings of the National Academy of Sciences*, *109*, 13526–13531.

Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kind? In G. Semin & K. Fiedler (Eds.), *Language and social cognition* (pp. 11–36). London: Sage.

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.

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*.

Taylor, M. G., Rhodes, M., & Gelman, S. A. (2009). Boys will be boys, cows will be cows: Children's essentialist reasoning about human gender and animal development. *Child Development*, *79*, 1270–1287.

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.