Child Development, September/October 2008, Volume 79, Number 5, Pages 1270–1287

Categories Influence Predictions About Individual Consistency

Marjorie Rhodes and Susan A. Gelman
University of Michigan

Predicting how people will behave in the future is a critical social-cognitive task. In four studies ($N = 150$, ages preschool to adult), young children (ages 4–5) used category information to guide their expectations about individual consistency. They predicted that psychological properties (preferences and fears) would remain consistent over time after hearing one example in which properties followed a category-linked distribution (e.g., children of different genders had different properties) but not when properties varied within a category (e.g., children of the same gender had different properties). The developmental course of these findings is examined. Results suggest the importance of considering how children's emerging theories of behavior and of social groups operate together to inform their expectations about the social world.

To an adult, hearing a child say that she ''likes Winnie the Pooh'' provides valuable information. For example, the adult may use this information to select birthday presents for this child or to choose books to offer to read to her. These inferences are possible because adults interpret the child's statement as stemming from a stable disposition toward Winnie the Pooh and expect this disposition to cause similar behaviors across time and situations (e.g., Yuill, 1993).

The example above demonstrates a tendency to explain human behavior by appealing to internal stable causal mechanisms, a lay theory termed *dispositionism* (Choi, Nisbett, & Norenzayan, 1999; Ross & Nisbett, 1991). Adults apply a disposition-based theory to explain and predict human behavior across a wide range of tasks, including when they have very limited information about the relevant behavior or characters (Gilbert & Malone, 1995; Uleman, 1987). A number of studies indicate, however, that preschoolers are less likely than adults to appeal to dispositions to explain behavior (Miller, 1984) or to expect individual behavior to remain stable across time and situations, especially when limited information is available (e.g., Kalish, 2002).

This research was supported by the National Institute of Child Health and Human Development Grant HD-36043 to the second author. The first author was supported by funding from the Michigan Prevention Research Training Grant (National Institutes of Health Grant T32 MH63057-03). We are grateful to the parents, teachers, and children of Lakewood Elementary School, Little Folks Corner, the Ann Arbor YMCA Early Childhood Center, and Concord Preschool for participating in this research, as well as to Korie Zink, Erica Ranade, and Kristin Garrison for their assistance with data collection. We also thank three anonymous reviewers who gave helpful feedback on prior drafts and whose suggestions motivated the design of Study 4.

Correspondence concerning this article should be addressed to Marjorie Rhodes, Department of Psychology, University of Michigan, 2422 East Hall, 530 Church Street, Ann Arbor, MI 48109-1043. Electronic mail may be sent to rhodesma@umich.edu.

The goal of the present article was to examine how preschoolers approach the task of explaining and predicting behavior when they have little evidence to draw on. If young children viewed most brief observations of behavior in their everyday lives as uninformative about the future, they would have some difficulty establishing a general sense of people and their roles, which would be a considerable impediment to learning from their environment. In the present work, we examined the possibility that preschoolers rely on category knowledge to make predictions about individual future behavior. In doing so, we aimed to enrich our understanding of young children's social inferences, by examining the circumstances in which they do and do not expect people to be consistent over time, as well as to examine a previously unexplored way that categories may guide induction in early childhood.

Dispositional Thinking in Early Childhood

Previous research from the related area of trait concepts suggests that dispositional thinking is accessible to preschool children under certain circumstances (e.g., Cain, Heyman, & Walker, 1997; Heyman & Gelman, 2000) but also that young children require more (or different kinds of) evidence than adults do to engage in such reasoning (Aloise, 1993; Boseovski & Lee, 2006; Miller & Aloise, 1989). Much of the research examining dispositional thinking has tested whether children and adults use information about past experiences to guide predictions about the future. For

example, researchers have examined whether learning that a girl liked chocolate cookies in the past leads participants to predict that she will like chocolate cookies again in the future (Kalish, 2002) or whether learning that a boy acted generously in one situation will lead participants to predict that he will act generously again in a different situation (Rholes & Ruble, 1984). On such questions, participants are often provided with relatively limited information about an actor's behavior (e.g., one example). Under these conditions, adults readily expect preferences and traits to be consistent over time, whereas children younger than age 7 demonstrate chance-level responding (Kalish, 2002; Kalish & Shiverick, 2004; Lawson & Kalish, 2006).

Using other methods, however, such as providing young children with many detailed examples of distinctive past behavior (Boseovski & Lee, 2006; Cain et al., 1997; Ferguson, van Roozendaal, & Rule, 1986; Gnepp & Chilamkurti, 1988), more sensitive response measures (Dozier, 1991), or linguistic cues (Gelman & Heyman, 1999), researchers have found that even 4-year-olds will predict behaviors and traits to be consistent over time. Also, a number of studies indicate that preschoolers can use trait terms to make predictions about behavior (Heyman & Gelman, 1999, 2000; Liu, Gelman, & Wellman, 2007) and that 3- and 4-year-olds can label examples of behavior with appropriate trait terms (Boseovski & Lee, 2006; Liu et al., 2007), indicating some facility with dispositional thinking when experimental tasks are simplified (but see Alvarez, Ruble, & Bolger, 2001; Yuill & Pearson, 1998, for alternate interpretations of preschoolers' success on tasks measuring dispositional thinking).

Kalish (2002) addressed preschoolers' hesitancy to predict consistency based on limited evidence, in spite of their knowledge of dispositions revealed on some tasks, with two proposals: (a) individuals predict consistency only when they view the event in question as resulting from a stable causal mechanism that is likely to cause the event again, and (b) preschoolers are less likely than adults to view behaviors as resulting from stable dispositions. That is, dispositions are not salient causal mechanisms within young children's theories of social behavior. Instead, Kalish suggests that young children are more likely to appeal to temporary factors to explain human behavior, such as passing desires and individual volition, and may not understand the role of stable dispositions in constraining those factors (see also Wellman, 1990). Thus, in the example given at the beginning of this article, preschoolers may conclude that the character said she liked Winnie the Pooh because she wanted to (e.g., felt like saying it), wanted Winnie the Pooh at that moment, or liked Winnie the Pooh for an external reason, such as a rule or adult request (Kalish & Shiverick, 2004; Karniol & Ross, 1976). In other words, they may not take this statement as evidence of a stable disposition toward Winnie the Pooh (see also Yuill, 1997). Because the character's preference is not viewed as resulting from a stable mechanism, preschoolers are unlikely to predict that the preference will be stable over time.

Within the framework described by Kalish (2002), it is not necessary to view dispositionism as entirely absent from young children's thinking; rather, it may be that young children require more clues before inferring the presence and importance of dispositions. For example, preschoolers may not view dispositions as relevant to as many behaviors as adults do or they may not view dispositions as common within human psychology. Thus, they would be less likely to appeal to dispositions to explain behavior if presented with little information. From this perspective, adults are more ready to predict consistency in social behavior because of the increased salience of dispositions within their naïve theories (e.g., Miller, 1984).

Category-Based Social Reasoning

We propose that when young children do not appeal to dispositions to predict behavior, they may incorporate their knowledge of other causal mechanisms into their social reasoning to make inferences. Preschoolers perceive membership in social categories, in gender categories in particular, as salient and powerful predictors of behavior. For example, Gelman, Collman, and Maccoby (1986) found that 4-year-olds expected two individuals of the same gender to share novel physical and behavioral properties, even when they were perceptually dissimilar. Similarly, Berndt and Heller (1986) found that 5-year-olds relied on gender categories to predict what sorts of activities characters would engage in (expecting them to engage in gender-typical more than gender-atypical behavior), even when they were told that characters had previously made gender-atypical behavioral choices (see also Biernat, 1991). In this case, kindergarteners relied more heavily on gender categories than adults did; adults attributed greater predictive power to characters' previous behaviors than to gender. Preschoolers also predict that children will display stereotypical gender behavior even if they are raised entirely by members of the other gender, whereas adults account for environmental influence in the development of gender-typical behavior (Taylor, 1996; Taylor, Rhodes, & Gelman, in press).

As discussed in detail by Lawson and Kalish (2006), preschoolers' willingness to base inferences on single pieces of evidence on category-based induction tasks, such as those involving gender, is in stark contrast to their hesitancy to use a single piece of evidence to make an inference about an individual over time. Preschoolers appear to consider membership in certain categories as predictive of a wide range of properties, such that they expect members of the same category to share many nonobvious properties, even if they differ from each other in other ways, such as physical appearance (for a review, see Gelman, 2003). Further, preschoolers expect membership in these informative categories to remain stable over time, regardless of environment (Gelman & Wellman, 1991; Hirschfeld, 1996; Taylor, 1996). Thus, whereas preschoolers may appeal to temporary factors to explain instances of individual behavior, they appeal to a stable factor (category membership) to make predictions about properties across category members.

We propose that in addition to appealing to category-based mechanisms to predict consistency across category members, preschoolers attend to social cues that guide them to appeal to categories to understand and predict instances of individual behavior as well. In situations where preschoolers do not perceive dispositions as causing behavior, they may attend to evidence indicating that the behavior is caused by membership in an important social category (e.g., gender). If such evidence is found, we expect preschoolers to infer that the behavior is caused by membership in a gender category, and therefore, to predict that the behavior will be consistent over time. Because gender categories are salient to young children, we expect that they will conclude that gender is causally relevant to behavior based on very limited information, as well as when the relation between gender and behavior is only implicitly communicated to them. Specifically, we hypothesize that preschoolers will predict consistency in individual behavior when they learn that individuals of one gender category, and not the other, display the properties of interest.

The Present Studies

In these studies, we first developed questions about characters' preferences and choices, in which participants received limited information about characters' previous experience. All questions involved characters engaging in novel activities, such that children's prior beliefs about the particular activities would not influence their pattern of responses. These questions did not elicit disposition-based reasoning among preschoolers, consistent with prior research. Next, we presented children with questions about pairs of characters who each experienced two things and had unique reactions to them. For example, a pair of children played two games; one child liked one game, whereas the other child liked the other. In some questions, the pairs involved two children of different genders, such that a boy liked one game and a girl liked the other. In this case, we expected children to perceive the pattern as indicating that liking a particular game is causally related to gender and, therefore, to expect preferences to be stable over time. The options were never labeled as ''for boys'' or ''for girls''; rather, children were given information about the preferences of individual male and female characters and were left to draw their own conclusions about the relation between gender and behavior. In other questions, participants were told about pairs of children of the same gender who each liked different activities. In this case, because the pattern does not imply that liking a particular activity is causally related to gender, we did not expect preschoolers to predict consistency over time.

Developmental Changes in the Use of a Category-Based Strategy

We also included older comparison groups in these experiments, in order to examine how the use of a category-based strategy to predict individual consistency changes over time. Because adults require less information than preschool children do to make dispositional inferences (Aloise, 1993), rely on a category-based strategy as often. On questions in which adults view characters' behaviors as relevant to dispositions, they should predict consistency regardless of category information.

Adults do not engage in dispositional thinking on all questions about individual behavior, however. For example, Kalish (2002) found that on questions asking whether characters would make consistent *choices* between two options that differed rather trivially (e.g., a boy who chose a blue flower over a yellow flower), adults responded randomly. This is likely because multiple strategies seem appropriate for answering this question; the boy might choose blue because of a stable disposition toward blue or he might choose yellow because he already has a blue flower (Kalish, 2002). Because no information about preferences is provided in these questions, both responses seem plausible.

We hypothesized that when adults perceive some ambiguity as to whether an actor's behavior should be considered as caused by a disposition, they too will turn to information about category membership to guide their predictions about consistency. We also hypothesized, however, that adults will appeal to category memberships only when they view the relevant categories as likely to have caused the property of interest. As reviewed above, adults view the influence of gender as less global and less deterministic than do younger children. Although adults often hold some strong gender stereotypes, in general they view gender as influencing a relatively narrower range of behaviors and view causal forces aside from gender (e.g., the environment, personality) as important predictors of behavior (Taylor, 1996). Therefore, given the paucity of information about the relation between gender and behavior presented in these questions, as well as that questions involve *novel* (i.e., nonstereotyped) behaviors, we expected that adults would be less likely than preschoolers to assume that characters' choices in these stories were caused by gender.

We included comparison questions about nonhuman animal species in order to determine whether any age-related changes in the use of a category-based strategy related to a complete shift away from viewing category-based mechanisms as relevant to individual behaviors or to developmental changes in the meaning of gender per se. These questions followed a parallel structure to all our questions about humans and gender, including questions about individual animals (e.g., one dog), pairs of same-species animals (e.g., two dogs), and pairs of different-species animals (e.g., a dog and a cat). We selected nonhuman animals because there is evidence that both young children and adults view animal species as indicating deep and fundamental differences between individuals and view many properties as causally related to membership in an animal kind (Gelman, 2003; Medin & Ortony, 1989). Therefore, we expected both preschoolers and adults to rely on species information to predict consistency (on questions in which they do not rely on dispositions).

Thus, young children were expected to reason similarly about gender categories and animal species, treating both kinds of categories as strongly influencing behavior, whereas adults were expected to show more differentiated reasoning about gender and animal species (Taylor, 1996; Taylor et al., in press). In later experiments, we also included a comparison sample of 10-year-olds to examine more precisely the development of differentiated reasoning about gender and species categories on these questions. Because we used novel properties in all questions, we were able to compare reasoning about humans and nonhuman animals directly. This research will therefore contribute to our understanding of how young children use category information to make predictions about individual consistency as well as of the influence of different kinds of categories on reasoning across development.

Overview

The aim of Study 1 was to develop a series of questions about single individuals and their psychological properties to be used in our subsequent experiments. In Study 2, we added information about category membership (of gender and animal categories) to the questions, such that preference for a particular activity either did or did not vary according to category membership. In Study 3, we expanded the type of properties and age ranges included. Finally, in Study 4, we demonstrated that preschoolers' predictions are influenced by gender categories but not by more superficial differences (i.e., shirt color).

Study 1

The goal of Study 1 was to develop questions to be used in our subsequent experiments. Participants were told about single individuals (animals or humans) who experienced two activities, had a particular response to them, and then a few days later were presented with the same two options again. We purposefully provided participants with limited information about prior behavior so that in subsequent experiments we could assess whether preschoolers engage in category-based reasoning to approach these questions. We included two types of psychological properties, preferences and choices, which we expected to differ in the extent to which they cued adults to appeal to dispositions. Thus, we hypothesized that preschoolers would demonstrate chance-level responding on all questions, whereas adults would predict consistency on questions about preferences but not on questions about choices.

Method

Participants. Participants (n = 30; 13 male and 17 female) included 14 preschoolers (M age = 4 years 8 months, range = 4 years 0 months – 5 years 7 months) recruited from private preschools and 16 adults (M age = 19 years 2 months, range = 17 years 8 months – 22 years 1 month) recruited from introductory

psychology classes at a large Midwestern public university. In all studies, the samples were approximately 80% White, 15% Asian American, and 5% from other racial/ethnic backgrounds. All participants were tested individually in a quiet office (located either at the children's schools or an on-campus laboratory) by the first author or by trained female undergraduate research assistants who were blind to study hypotheses. College students received course credit for their participation.

Procedure. Participants were presented with 16 questions, including 8 questions about children (4 about boys and 4 about girls) and 8 questions about animals of various species. Half of the questions asked about characters' preferences and half asked about their choices. Thus, there were 4 questions of each of four types (e.g., human preferences, human choices, animal preferences, and animal choices). All questions followed an identical format (see Appendix A).

For example, on a question about preferences, participants were told, "Here is Tom. A few days ago, Tom did two things: Tom went flecking and Tom went tooping. Tom liked flecking. Today, Tom is doing two things again. Again, Tom is going flecking and tooping. What will Tom like today? Will Tom like flecking, like last time, or will Tom like tooping?" After one question of each type, randomly selected in advance, participants were asked, "Remind me, which did [Tom] like the first time?" This allowed us to assess whether participants were following the scenarios; no adult and only 1 child had difficulty answering any of these questions. This child missed only one control question (of four total asked) and so was retained for analyses. All questions were accompanied by simple pictures of the story character as well as by small cards with the names of the activities printed on them. Pointing was used to aid comprehension. For all questions, responses in which participants predicted that the character would act in a consistent manner were scored as 1 (the alternate response was scored as 0).

Questions were blocked by character type (animal or human child), and order of presentation of the two blocks was counterbalanced across participants. Within each block, questions were presented in a separate random order for each participant. Several precautions were taken to prevent the results from being biased by any unforeseen preferences that the participants might have had for particular nonsense words. Specifically, the order of presentation of the two activities and the selection of particular novel activities by the story character were counterbalanced for each question across participants (e.g., in the example question above, half of the participants were told that Tom liked "flecking," whereas the other half were told that he liked "tooping"). Also, pairs of nonsense words circulated throughout the question types across participants, such that each pair of words was used equally in each of the four types of questions.

Results

We first compared the proportion of questions on which participants predicted that characters would behave consistently to the proportion expected by chance (chance = .5) for each type of question, using a series of sign tests. As shown in Figure 1, on questions about preferences, preschoolers demonstrated chance-level responding (human children: $M = 0.48$, $SD = 0.37$; animals: $M = 0.44$, $SD = 0.65$), whereas adults predicted consistency more often than expected by chance (human children: $M = 0.78$, $SD = 0.22$; animals: $M = 0.83$, $SD = 0.22$). On questions about choices, preschoolers again did not predict consistency; they demonstrated chance-level responding on questions about animals ($M = 0.37$, $SD = 0.42$) and predicted consistency *less* often than expected by chance on questions about human children ($M = 0.25$, $SD = 0.35$). As hypothesized, adults also did not reliably predict consistency on questions about choices (human children: $M = 0.41$, $SD = 0.31$; animals:

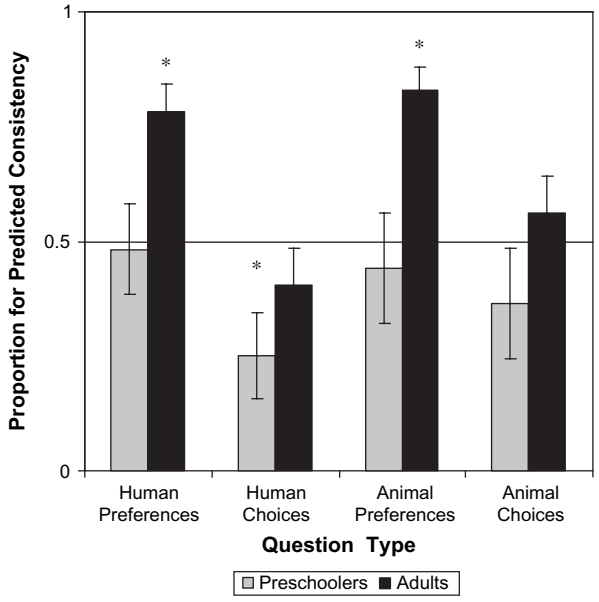

Figure 1. Proportion for predicted consistency by age and question type, Study 1.
Note. Error bars represent standard errors of the mean. Sign tests comparing the obtained proportions of predicted consistency to the proportion expected by chance.
*$p < .05$.

$M = 0.56$, $SD = 0.34$). There were no main or interactive effects of participant gender or character gender, so these variables were excluded from analyses.

We next examined the data through a 2 (age: preschooler, adult) × 2 (character: human, animal) × 2 (question type: preference, choice) repeated measures analysis of variance, (ANOVA), with character and question type as within-subjects variables and age as a between-subjects variable. This analysis yielded main effects for age, $F(1, 27) = 6.12$, $p < .05$, such that adults predicted consistency more often than did preschoolers, and for question type, $F(1, 27) = 24.83$, $p < .001$, such that consistency was predicted more often for preferences than for choices. There were no other significant main or interactive effects. The interaction between question type and age was marginal, $F(1, 27) = 3.06$, $p = .09$; however, follow-up analyses indicated that the effect of age was specific to questions about preferences. As demonstrated in Figure 1, adults predicted consistency more often than preschoolers on questions about preferences, for humans, $t(28) = 2.71$, $p < .05$, $d = .99$; for animals, $t(28) = 3.0$, $p < .01$, $d = 1.10$, but not on questions about choices, for humans, $t(28) = 1.28$, *ns*; for animals, $t(28) = 1.41$, *ns*.

Discussion

The results from Study 1 support the hypothesis that preschoolers are less likely than adults to predict that preferences are consistent over time; adults strongly expected preferences to be consistent over time, whereas children responded at chance. Also, as predicted, adults did not reliably predict consistency on questions about choices. Interestingly, on the questions about human choices presented in this study as well as on some similar questions about choices asked by Kalish (2002), young children actually chose the alternate (inconsistent) response more often than expected by chance. This suggests that children may expect others to favor variety *over* consistency, especially when no preference information is provided.

Although these findings regarding children's beliefs about choices are interesting, in Study 2, we decided to ask children questions about preferences only, for two reasons. First, the aim of this series of studies was to examine whether children would use category information to guide their predictions about consistency on questions on which they otherwise would not respond in an adult-like manner, and adults reliably expected consistency only on questions about preferences. Second, pilot testing demonstrated that because the questions asked in Study 2 contained more information than those from Study 1 (see below), the task was too long for young children when questions about both preferences and choices were asked.

Based on these considerations, both preschoolers and adults in Study 2 were asked questions about preferences. For adults only, we also included supplemental questions about choices. These were included because adults' predictions of consistency on questions about preferences were relatively high, indicating that they were engaging in dispositional thinking on these questions. Therefore, we did not expect information about category memberships to influence their predictions. Results from Study 1 indicated, however, that they did not engage in dispositional thinking on questions about choices; therefore, adding category information to these questions allowed us to test the hypothesis that adults are influenced by information about animal species, and not gender categories, on these questions.

Study 2

In Study 2, we tested whether guiding participants to view individuals' psychological properties as linked to category membership would lead them to predict greater levels of individual consistency. Participants were asked questions about pairs of either animals or human children that followed a similar format and referred to the same activities as those asked in Study 1. In Study 2, however, participants were given scenarios about two characters in which one character preferred one activity and the other character preferred the other activity. We hypothesized that when the pairs consisted of two individuals that shared category membership (e.g., two children of the same gender or two animals of the same species), the property would not be viewed as linked to category membership (as two members of the same category have two different preferences), and therefore, young children would again fail to predict consistency. In contrast, when participants are told about pairs consisting of members of different categories (e.g., two children of different genders or two animals of different species), we predicted that children would then view the property as category linked and would predict greater levels of individual consistency. We hypothesized that adults would readily predict consistency for all questions about preferences, regardless of category information. On the supplemental questions about choices, however, we hypothesized that they would be influenced by information about pair composition on questions about animal species but not human gender.

Method

Participants. Participants (n = 53, 23 male and 30 female) included 24 preschoolers (M age = 4 years 7 months; range = 3 years 11 months – 5 years 8 months) and 29 adults (M age = 19 years 1 month; range = 17 years 9 months – 22 years 7 months) drawn from the same population as those who participated in Study 1; no one participated in both studies.

Procedure. All participants were tested individually in a quiet office (located at either the children's schools or an on-campus laboratory) by the first author or by trained female undergraduate research assistants who were blind to study hypotheses. College students received course credit for their participation.

To limit the number of questions that each child was asked, participants were randomly assigned to receive questions about either animals or humans. Each condition included eight questions, which consisted of two blocks of four questions each, one block asked about pairs of individuals from the same category (i.e., same gender or same species) and the other block asked about pairs of individuals from different categories (i.e., different gender or different species). Order of presentation of the two blocks was counterbalanced across participants, and the questions within each block were asked in a separate random order for each participant. For adults, each block contained an additional four questions describing choices instead of preferences (see Appendix B). The same procedures as in Study 1 were used for assigning nonsense words to questions for each participant and for counterbalancing the order of nonsense words within each pair.

Questions followed the same structure and used the same properties as in Study 1, except that pairs of characters were described and none were given individual names (see Appendix B). For example, children heard: "Here are two children, a boy and a girl. A few days ago these children did two things, they went lexing and they went gorping. The boy liked lexing. The girl liked gorping. Today, the boy is doing those two things again. Again, the boy is going lexing and gorping. What will the boy like today? Will the boy like lexing, like last time, or will the boy like gorping?" After one question of each type, randomly selected in advance, participants were asked to recall the details of the story after they made their predictions, with the following questions, "Remind me ... which one did [the boy] like the first time? And, which one did [the girl] like?" No participant had difficulty answering these control questions. Questions were accompanied by pictures of both characters and cards with the names of the activities printed on them. Pointing was used to aid comprehension. After the first set of preferences was described, the picture of the nontarget character was removed.

Results

We first present the analyses examining children's and adults' predictions about preferences, followed by supplemental analyses examining adults' predictions about choices. Preliminary analyses revealed that there were no main or interactive effects of participant gender or character gender; therefore, these variables were not included in analyses.

As shown in Figure 2, sign tests revealed that preschoolers predicted consistency at levels exceeding those expected by chance for questions about pairs of individuals from different categories (different gender: M = 0.71, SD = 0.32; different species: M = 0.85, SD = 0.27). In contrast, preschoolers predicted consistency less often than expected by chance on questions about same-gender pairs (M = 0.27, SD = 0.33) and at chance levels for same-species animal pairs (M = 0.60, SD = 0.34). Adults predicted consistency more often than expected by chance for all questions about preferences (same gender: M = 0.69, SD = 0.28; different gender: M = 0.75, SD = 0.27; same species: M = 0.83, SD = 0.28; different species: M = 0.89, SD = 0.23).

We further examined these data through a 2 (age: preschooler, adult) × 2 (condition: human

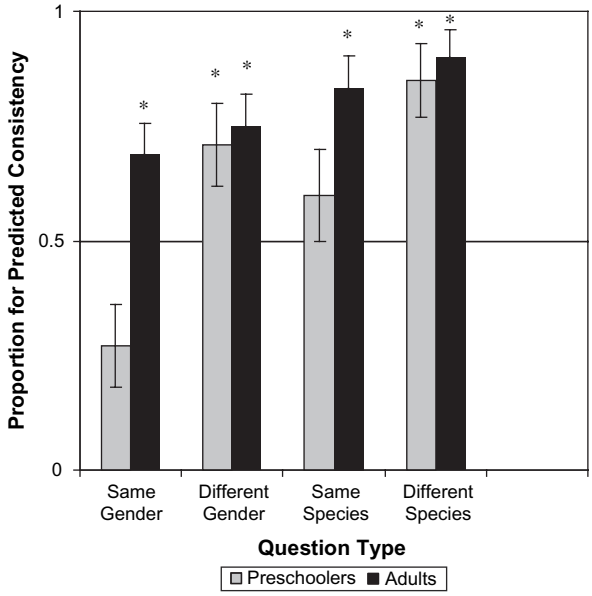

Figure 2. Proportion for predicted consistency by age and pair composition, Study 2.
Note. Error bars represent standard errors of the mean. Sign tests comparing the obtained proportions of predicted consistency to the proportion expected by chance.
*$p < .05$.

children, animals) × 2 (pair composition: same category, different category) repeated measures ANOVA, with age and condition as between-subjects factors and pair composition as a within-subject factor. This analysis revealed main effects for pair composition, $F(1, 49) = 18.15$, $p < .001$, such that more consistency was predicted on questions about members of different categories; condition, $F(1, 49) = 8.86$, $p < .01$, such that more consistency was predicted for animals than people; and age, $F(1, 49) = 8.29$, $p < .001$, such that adults predicted consistency more often than did children. There was also a significant Pair Composition × Age interaction, $F(1, 49) = 8.48$, $p < .01$; see Figure 2. Planned contrasts revealed that preschoolers predicted consistency more often on questions involving pairs from different categories than on questions involving pairs from the same category ($p < .001$, $d = 1.20$), but adults' predictions did not differ depending on the composition of the pairs ($p > .5$). This pattern was identical and significant for questions about human and animal targets. Also, for same-category pairs, preschoolers predicted consistency less often than adults did ($p < .001$, $d = -1.04$), but for different-category pairs, they predicted consistency as often as did adults ($p > .3$).

Adults' predictions about choices. As hypothesized, adults predicted consistency more often than expected by chance only for questions describing animals of different species ($M = 0.75$, $SD = 0.28$, $p < .01$, sign test) and not for any other question type (same species: $M = 0.50$, $SD = 0.45$; different gender: $M = 0.45$, $SD = 0.26$; same gender: $M = 0.45$, $SD = 0.31$). Also as hypothesized, adults' predictions on questions about animals differed depending on the composition of the pair; adults predicted consistency more often on items describing different-species animals than on items describing same-species animals, $t(14) = 2.49$, $p < .05$, $d = .66$. In contrast, adults' predictions on questions about humans did not differ depending on the gender composition of the pair, $t(15) < .001$, *ns*

Discussion

In Study 2, we found support for the hypothesis that guiding preschoolers toward viewing individuals' preferences as linked to category membership leads them to predict greater individual consistency. Children predicted that preferences would remain consistent after learning that members of different categories had different preferences but not after learning that members of the same category had different preferences. We also documented that whereas adults' predictions about properties for which they have strong expectations of consistency (e.g., preferences) are not influenced by category information, they do attend to category information on questions about properties for which their expectations of consistency are less certain. Adults' concepts of different kinds of categories also appear to influence whether they view a category as relevant to individual consistency on these questions; adults were influenced only by information about animal species not by information about gender categories.

The results from adults' predictions about choices provide preliminary support for our hypothesis that categories are only used to predict individual consistency when the category is viewed as causally relevant to novel behaviors. We have not yet been able to document age-related changes in the use of gender and animal species on questions about the same psychological property, however, because adults predict preferences to be consistent so readily, and young children in this study were not asked questions about choices. We address this limitation in Study 3 by including a comparison group of 10-year-olds. We selected 10-year-olds for two reasons. First, previous research indicates that 10-year-olds predict consistency more often than do younger children, but we did not expect their predictions to be as stable as those of adults; therefore, we expected their predictions to be somewhat influenced by information about categories. Second, previous research indicates that by age 10, children have substantially refined their concepts of gender and view the influence of gender as relatively probabilistic and specific, as opposed to as deterministic and global (e.g., Berndt & Heller, 1986; Taylor, 1996; Taylor et al., in press). Therefore, we expected their inferences—like those of adults—to be influenced by information about animal species but not human gender. Comparing the responses of 10-year-olds with those of younger children will allow for a cleaner test of the hypothesis that beliefs about particular categories influence whether children appeal to category-based mechanisms.

Study 3

In Study 3, we further examined the developmental course of the results presented in Studies 1 and 2 by including a group of 10-year-olds. To maximize comparability between the two age groups, our younger participants in this study were drawn from kindergarten classes from the same elementary school attended by the 10-year-olds, who were drawn from fifth-grade classrooms. Thus, the participants in the younger age group of this study were slightly older than in Studies 1 and 2 but still within the age range

of young children who are unlikely to engage in dispositional thinking on questions that provide limited evidence (e.g., Kalish, 2002).

We also made two small modifications designed to address the generality of our findings. First, we provided participants with information about what kinds of activities the characters engaged in. In Studies 1 and 2, participants were told relatively little information, for example, that characters had "gone lexing." They were left to infer that "lexing" is some sort of activity, but it could be a game, a chore, or any other type of activity. In Study 3, we examined whether children would use the same category-based strategies as they had in Study 2 when information about the types of activities that characters engaged in is explicitly given. For example, in these questions, children are told about boys and girls who played two games or bears and monkeys that climbed two kinds of trees. The options were still labeled using nonsense words (e.g., to refer the particular games or trees) in order to control for the influence of children's prior knowledge or experiences (e.g., their preferences for specific games). For example, children were told that characters played the "lex game and the gorp game."

Second, in Study 3, we expanded the types of psychological properties examined. In Study 2, children were asked questions about preferences only. Because preferences are an area that children frequently think about in terms of gender (e.g., games that girls like), it is possible that children's gender-based strategy is limited to questions about preferences. Alternately, because young children view gender as having a broad influence on psychological properties, they may apply this strategy more liberally. To distinguish between these possibilities, in Study 3, we included questions about both preferences and fears. We selected fears because older children tend to view fears as stable over time, whereas younger children may not (e.g., Kalish, 2002; Yuill, 1997).

Because Study 3 involved new sets of questions and new ages, we first conducted an initial pilot study following the same format as in Study 1. The goal of this pilot study was to confirm that 10-year-olds predicted consistency more often than 5-year-olds on questions about single individuals, using the new properties created for this study, when no information about categories is provided. After the pilot study demonstrated the expected age-related differences on our new items, we conducted the experimental study, in which participants were asked a series of questions about pairs of children and animals (similar to Study 2).

Method

Participants. Participants (n = 55; 26 male and 29 female) included 25 kindergarteners (M age = 5 years 7 months, range = 4 years 9 months – 6 years 5 months) and 30 fifth graders (M age = 10 years 1 months, range = 9 years 2 months – 11 years 4 months) recruited from a single public elementary school in a Midwestern university town. Eleven of the kindergartners and 15 of the fifth graders were assigned to the pilot study, and the rest were assigned to the main study (described below).

Procedures. Participants were tested in a quiet office at their elementary school by either the first author or trained female undergraduate research assistants who were blind to study hypotheses. In both the pilot study and the main study, participants were asked a total of 16 questions (8 about animal pairs and 8 about human pairs). Questions were blocked by character type, and order of presentation of these blocks was counterbalanced across participants; kindergarteners completed the blocks in two sessions, approximately 1 week apart; fifth graders completed the study in a single session. Within each block, questions were also blocked by pair type (same category or different category), with order of these blocks also counterbalanced across participants. Within each block, questions were presented in a separate random order for each participant. Procedures for counterbalancing nonsense words were the same as in previous studies.

Pilot study. In the pilot study, participants were asked questions about individual actors—eight human characters and eight animal characters—which followed the same structure as the questions used in Study 1 but referred to the new properties developed for this study. Half of the questions asked about preferences and half asked about fears, such that there were four questions of each of four types (four animal preferences, four animal fears, four human preferences, and four human fears; see Appendix C). Results from this pilot study indicated that the questions using the new properties elicited the expected age-related findings. Specifically, kindergarteners did not predict consistency more often than expected by chance on any questions (human children: M = 0.36, SD = 0.44; animals: M = 0.30, SD = 0.43; $p > .34$, sign tests), whereas fifth graders predicted consistency more often than expected by chance for both kinds of characters (human children: M = 0.70, SD = 0.33; animals: M = 0.70, SD = 0.31; $p < .05$, sign tests). Also, fifth graders predicted consistency more often than did kindergarteners on questions about both humans, $t(24) = 2.22$, $p < .05$, $d = .87$, and animals, $t(24) = 2.78$, $p < .05$,

d = 1.08. There were no effects associated with property type (preference or fear).

The questions asked in this pilot study were similar in design and purpose to those asked in Study 1, with three changes: (a) the properties were modified as described above, (b) the younger participants were slightly older than in Study 1, and (c) category labels were used to refer to the characters in the question (e.g., "This is a girl") instead of proper names (e.g., "This is Molly"). Despite these changes, the proportion of questions on which kindergarteners predicted consistency did not differ significantly from the proportion of questions on which preschoolers in Study 1 predicted consistency, on questions about human children, $t(23) = .73$, *ns*, or animals, $t(23) = .82$, *ns*. Based on these findings, we developed the questions for the main study using the new property descriptions.

Main study. In the main study, participants were asked eight questions about animals (four same-species pairs and four different-species pairs) and eight questions about children (four same-gender pairs and four different-gender pairs). Thus, there were four questions of each of four types (see Appendix D). Procedures were similar to Study 2 and materials included the same pictures of the story characters, as well as cards containing the names of the target activities.

Results

In the main study, we first compared the proportion of questions on which participants predicted that characters would behave consistently to the number expected by chance for each type of question, followed by repeated measures ANOVAs. Preliminary analyses revealed that there were no main or interactive effects of participant gender or character gender; therefore, this variable was not considered in analyses. Also, there were no differences between preferences and fears for either age group; in no analysis did the main or interactive effect of property type (preference or fear) approach significance (all ps $> .4$). Therefore, we collapsed across this distinction.

As shown in Figure 3, sign tests revealed that kindergarteners predicted consistency more often than expected by chance on items involving different category members for questions about both human gender (M = 0.80, SD = 0.33) and animals (M = 0.73, SD = 0.40) but not for questions involving same-category pairs (same gender: M = 0.41, SD = 0.35; same species: M = 0.43, SD = 0.36). Fifth graders predicted consistency more often than expected by chance for items involving different animal species only (M = 0.75, SD = 0.27) and not for items involving same-species animals (M = 0.63, SD = 0.39), same-gender children (M = 0.68, SD = 0.35), or different-gender children (M = 0.62, SD = 0.38).

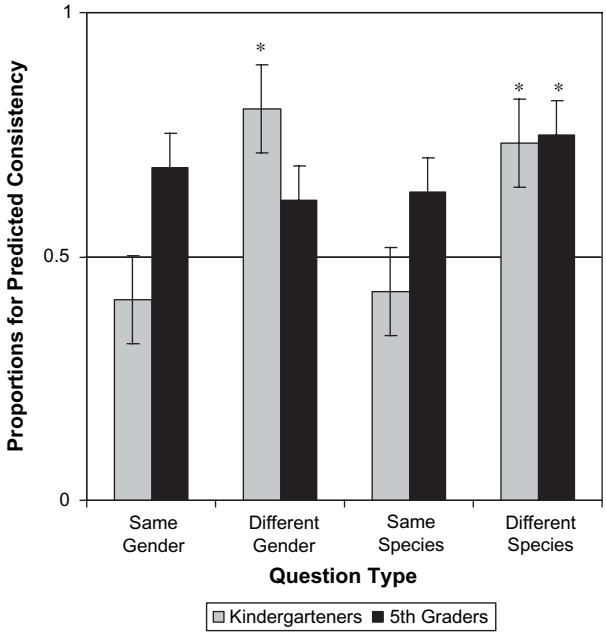

Figure 3. Proportion for predicted consistency by age and pair composition, Study 3.
Note. Error bars represent standard errors of the mean. Sign tests comparing the obtained proportions of predicted consistency to the proportion expected by chance.
*$p < .05$.

We further examined these data through a 2 (target: animal, human) × 2 (pair composition: same category, different category) × 2 (age: kindergarten, fifth grade) repeated measures ANOVA, with target and pair as within-subjects factors and age as a between-subjects factor (see Figure 3). This analysis revealed a main effect of pair composition, $F(1, 27) = 16.64$, $p < .001$, and a significant interaction between pair composition and age, $F(1, 27) = 12.48$, $p < .01$. Planned contrasts revealed that kindergartners predicted consistency more often on questions about different-category members than on questions about same-category members, for both human and animal characters ($p < .001$; gender, d = 1.15; animal, d = .79). Although the three-way interaction among target, pair composition, and age failed to reach significance ($p = .1$), for fifth graders, as hypothesized, pair composition did not influence responses on items about gender: Predictions about same-gender and different-gender pairs did not differ ($p > .3$). However, pair composition did influence responses on items about animals: Predictions of consistency were made significantly more often about different-species pairs than about same-species pairs, $t(14) = 2.43$, $p < .05$, d = .35.

Discussion

In Study 3, we replicated the findings from Studies 1 and 2 for younger children and extended these findings to questions about fears. When presented with questions involving individual human children or animals, kindergartners did not expect the characters to have consistent preferences or fears over time (as demonstrated in the pilot study). When presented with pairs of children or animals, however, kindergartners expected characters to have consistent preferences and fears only when they were told that members of different categories had different properties.

In contrast, fifth graders generally expected humans and animals to have consistent preferences and fears, as evidenced in the pilot study. Also as hypothesized, fifth graders predicted consistency more often on questions about different-species animals than on questions about same-species animals but did not differentiate between questions about same-gender and different-gender pairs. These findings are consistent with prior work demonstrating that older children view gender as less influential for behavior than do younger children (e.g., Berndt & Heller, 1986; Taylor, 1996). These findings, as well as the findings from Study 2, suggest that developmental changes in how gender influences predictions of consistency on these questions relate to changes in children's gender concepts not to a general move away from viewing category information as relevant to individual consistency.

Surprisingly, fifth graders demonstrated chance-level responding on all questions in the main study, except for questions involving pairs of different animals. Thus, learning that pairs of children had different preferences and fears may have led fifth graders away from predicting consistency. Although the predictions of fifth graders did not differ from chance in the main study, they did still predict consistency more often than kindergarteners. Nevertheless, it is somewhat surprising that fifth graders' predictions on these questions were not above chance, as they are well within the age range of children that have been shown to engage in dispositional reasoning in other studies (e.g., Kalish, 2002).

Children's spontaneous explanations that accompanied these questions provide some insight into this unexpected pattern. Participants frequently mentioned various social learning mechanisms in explaining their responses. For example, children suggested that characters might ''learn to like it from their friend'' or ''become afraid of it because their partner did'' (explanations fitting this interpretation were given at least once by 8 of 14 fifth graders in this condition). Thus, children hypothesized that social interactions between the characters might influence the tendency to try new things or remain consistent. Importantly, this process was not influenced by the gender composition of the pair, supporting the hypothesis that fifth graders would not view gender as influencing the consistency of these psychological properties. It is also possible, given children's explanations, that their beliefs about the social context presented in the study questions somehow guided them away from viewing gender as relevant. Species information, however, did influence this pattern, as described above. Thus, although children posited social interactions that would lead the pairs of children and animals to influence each other, they did not expect such interactions to operate across species boundaries (although they did expect them to operate across gender categories). Overall, these findings support the prediction that, with age, children develop more differentiated concepts of animal species and gender categories.

Study 4

Study 4 was a control study, which was designed to test whether young children appeal to category-based mechanisms only when the category information that is presented refers to meaningful differences between people. In Studies 2 and 3, young children reliably predicted consistency on questions involving both different-gender humans and different-species animals. One possibility is that young children will appeal to *any* observable difference between actors to make predictions about consistency, even if the distinction is relatively unlikely to influence the behavior of interest. From this perspective, the more children are guided to focus on the differences between two characters, the more likely they may be to infer stable individual differences. If this were the case, then young children's responses on our previous studies would not be indicative of their beliefs about the causal power of gender or species categories per se and would not represent a theory-driven category-based approach to social reasoning. Alternately, if children use gender categories, but not relatively more trivial differences between people, to predict consistency, then the pattern of findings documented in our earlier studies may be interpreted as reflecting the conceptual status of gender categories and animal species. To distinguish between these two possibilities, we compared children's reasoning about pairs of different-gender children and pairs of children who differ on another easily observable (but less

meaningful) dimension, in this case, shirt color. We selected shirt color because it is easy for young children to track visually, amenable to labeling, and has been used in other studies designed to compare reasoning about deep and relatively superficial categories (Solomon, 2002).

Method

Participants. Participants included 12 preschoolers (5 male and 7 female; *M* age = 4 years 8 months, range = 4 years 1 month – 5 years 7 months) recruited from the same population areas as children who participated in Studies 1 and 2. None had participated in an earlier study.

Procedure. Children were tested individually, following the procedures used in previous studies. All children were asked a series of eight questions: four about different-gender pairs and four about pairs of children (of the same gender) who wore different-color shirts. Items were blocked by pair type (gender or shirt color), and order of presentation of the blocks was counterbalanced across participants. Within the blocks, items were presented in a separate random order for each participant. The questions about gender began, ''See these two children? This one is a boy. This one is a girl,'' and then were identical to the different-gender pair questions from Study 2. The questions about shirt color began ''See these two children? This one is wearing a yellow shirt. This one is wearing a green shirt,'' and then followed the same structure as the other questions (see Appendix E). As in previous studies, children were presented with simple pictures of the story characters, as well as with small cards with the names of the nonsense activities printed on them, and pointing was used to aid comprehension. New pictures were used for this experiment, so that we could control shirt color (for different-gender pairs, the children wore the same-color shirt; for different-shirt-color pairs, children were of the same gender, but one wore a yellow shirt and one wore a green shirt). As in previous studies, order of presentation of each nonsense word within a pair and the character's selected preference were counterbalanced across participants and the pairs of nonsense words circulated across items.

Results

Consistent with previous studies, preschoolers predicted that characters' preferences would remain consistent more often than expected by chance when they were told about different-gender pairs ($M = 0.73$, $SD = 0.31$, $p < .05$, sign test). As hypothesized, however, they did not predict that preferences would be consistent when they were told about pairs of children wearing different-color shirts; in fact, they predicted that preferences would be consistent marginally less often than expected by chance ($M = 0.27$, $SD = 0.39$, $p = .09$, sign test), which was similar to preschoolers' performance on questions about same-gender pairs in Study 2. Children predicted consistency more often on questions about gender pairs than on questions about shirt-color pairs, $t(11) = 5.01$, $p < .001$, $d = 1.53$.

Discussion

In this control study, we documented that preschoolers apply a category-based strategy only when the individual differences that are presented relate to meaningful categories (i.e., gender) but not when they relate to more trivial differences (i.e., shirt color). These findings support our interpretation that individuals appeal only to categories that they view as having an important causal influence on behavior to predict consistency and suggest that preschoolers appeal to categories in a strategic manner to predict the future.

General Discussion

We examined whether individuals rely on category information to predict individual behavior. We found that preschoolers, older children, and adults all rely on such a process to guide their social reasoning. We also found, however, that there are age differences in how commonly participants engaged in category-based reasoning, as well as in which kinds of categories individuals found informative for reasoning about these items.

We found robust evidence for a category-based approach to reasoning about individual consistency among preschoolers and kindergarteners. When young children were told that a single individual experienced two options and had a preference for or was afraid of one of them, they were no more likely to predict that those responses would be consistent over time than expected by chance, congruent with prior work (e.g., Kalish, 2002). When young children were told that two characters from different salient categories had unique responses (e.g., a boy liked one option, whereas a girl liked the other), however, they predicted that psychological responses would be consistent over time more often than expected by chance and as often as adults did. We found evidence for increased predictions of consistency only when the two characters differed in terms of membership in an important category; young children did not predict consistency when they were told about the unique

responses of two characters that shared category membership (e.g., two girls), even if the two characters differed in some other observable manner (e.g., shirt color).

Young Children's Category-Based Reasoning

Why do preschoolers predict consistency when they learn that members of different categories have different properties? It is well documented that young children view membership in certain categories as fundamental to identity and as determining a wide range of behaviors (Gelman, 2003). Much of this research has focused on children's beliefs about animal species and has demonstrated that young children make a broad range of inferences about animals' properties based on species information alone (Gelman & Markman, 1986; Gelman & Wellman, 1991; Sousa, Atran, & Medin, 2002; Waxman, Medin, & Ross, 2007). As described by Gelman, (2003), species categories take on this inductive potential because children view category membership as tied to the presence of a category ''essence,'' or a non-obvious internal entity that exists in every category member and causes category membership and the development of category-linked properties. Thus, the development of category properties is viewed as causally related to category membership.

A number of studies indicate that children also perceive membership in some social categories as resulting from the presence of a category essence (Astuti, Solomon, & Carey, 2004; Diesendruck & haLevi, 2006; Hirschfeld, 1996; Taylor, 1996). Preschoolers treat gender categories in particular as having high inductive potential and as causally determining behavior. For example, preschoolers expect members of the same-gender category to share novel properties even in the face of contradictory perceptual or environmental information (Gelman et al., 1986; Taylor, 1996). Thus, preschoolers appear to use gender categories and animal species information in a similar manner in inductive reasoning, treating them both as indicating membership in natural kinds (Rothbart & Taylor, 1992).

Information about category membership, whether regarding animal species or human gender, presents young children with powerful causal explanations for behavior. When confronted with the question of whether a particular behavior is likely to be consistent over time, the belief that a behavior is causally related to category membership provides a strong reason to predict consistency. We propose that because categories provide such important explanatory information, young children are sensitive to detecting category-linked distributions. In the present studies, they formed expectations that a behavior was tied to gender after a single piece of evidence. Behaviors were never labeled as ''for boys'' or ''for girls,'' children made this inference on the basis of information about only two individuals.

Developmental Changes in Category-Based Reasoning

Adults and older children (fifth grade) also relied on information about categories to guide their inferences about consistency, but only on questions involving animal species, and for adults, only on questions about particular psychological properties. We will discuss two aspects of the development of social cognition that we believe account for these age-related findings, one related to the increased salience of dispositionism as a lay theory for explaining behavior, and the other related to changing beliefs about gender.

Dispositionism. Kalish (2002) suggests that dispositions, or internal stable individual differences, are not salient causal mechanisms within young children's theories of behavior. Although previous work suggests that preschoolers have some capacity to engage in dispositional thinking (e.g., Heyman & Gelman, 2000), on problems about individual consistency they appear to require more or different kinds of evidence than older children and adults before they will view a given behavior as good evidence on which to base an inference of future consistency.

This interpretation is consistent with findings from cross-cultural research indicating increasing cultural divergence in social attribution with age. These cross-cultural findings suggest that social attribution is, at least in part, the result of an extended process of socialization (Miller, 1984, 1986). Wellman and Miller (2006) summarize this body of literature and suggest that although there are some cross-cultural commonalities in young children's understanding of people and behavior (e.g., viewing people as active and intentional agents), the specifics of children's naïve theories of psychology, including which factors they view as salient and powerful behavioral motivators, vary cross culturally. Among adults as well, some aspects of social attribution appear to be common across cultures (e.g., viewing individuals as active in initiating their own behavior), but individuals in different cultural contexts appear to differ in the extent to which they also attribute a potent causal role to the environment (Choi et al., 1999; Morris & Peng, 1994; Norenzayan, Choi, & Nisbett, 2002). Thus, attributing a primary causal role to individual dispositions appears to be the result of a process of

socialization, interacting with a probably universal tendency to view humans as intentional. From this perspective, it is not expected that preschoolers would share the same conceptual framework for explaining and predicting behavior as the adults in their environment, as much of socialization is yet to be completed (Wellman & Miller, 2006). Thus, the increasing salience of dispositionism with age in the present context may contribute to our finding that older children and adults relied on a category-based strategy less often than did younger children.

Gender concepts. A second developmental process that may contribute to the present findings relates to developmental changes in concepts of different kinds of categories. Young children reasoned about animal species and human gender in a similar manner, regardless of whether this variable was manipulated within or between participants, suggesting that young children attribute a similar causal role to both kinds of categories. Young children's relatively strict concepts of gender have been suggested to relate to a general tendency to focus on within-category similarity and between-category differences and to ignore within-category variability (Gelman, 2003; Gelman & Coley, 1991; Maccoby, 1988). Their approach to gender may also relate to specific beliefs about *how* gender influences behavior. For example, previous work suggests that preschoolers view engaging in gender-stereotyped behaviors as somewhat automatic, or as necessary for fulfilling moral or social obligations, as opposed to as subject to environmental or intentional control (Levy, Taylor, & Gelman, 1995; Taylor et al., in press).

In contrast, fifth graders' and adults' reasoning appeared to reflect more differentiated concepts of animal species and human gender; they were influenced by category information about animal species only. Prior work has suggested that, over time, individuals come to view gender as influencing a narrower set of properties, as well as to view the influence of gender on behavior as probabilistic and dependent on environmental pressure and personal agency (Berndt & Heller, 1986; Taylor, 1996). Although adults and older children often hold some strict stereotypes about gender, these may be more specific than the theories held by young children. With age, individuals' concepts of gender focus on differences in particular areas, whereas younger children expect more general differences (Maccoby, 1988). Thus, we propose that developmental changes in theories about gender also contribute to our age-related findings.

We have indicated that because young children view gender categories as such strong predictors of behavior, they are willing to infer a gender-linked cause for behavior after one example of a gender-linked pattern. In contrast, the presented evidence was not sufficient to warrant such an interpretation by older children or adults. We suspect, however, that older children and adults would rely on gender if they were presented with a greater amount of gender-linked information (e.g., questions involving stereotypical properties or more examples of gender-linked patterns). In future research, it would be interesting to compare how much gender-linked information preschoolers, older children, and adults require before they view novel behaviors as related to gender.

Limitations and Future Directions

A direction for future research suggested by these findings stems from the fact that the information about past experiences provided to participants in the present experiments was relatively minimal, as was the information provided about categories. In the present work, we found that when both of these sources of information are minimal (e.g., one example of a prior behavior and one example of a category-linked distribution), preschoolers relied on category-linked information but not dispositions to guide their inferences. One limitation of this approach, however, is that all the activities referred to in the present research were novel. Although using novel activities in this context was useful for experimental control, it is certainly possible that preschoolers will infer dispositions based on single acts if they are able to apply their knowledge about particular activities (e.g., they may view some activities, such as liking brussel sprouts, as distinct enough as to warrant dispositional reasoning).

Also, because the information provided was limited, we cannot determine from the present research when children might rely on a category-based strategy in more realistic contexts. Based on prior work, we suspect that if more information about previous behavior was provided, preschoolers may be more likely to view the behavior as linked to a disposition (e.g., Boseovski & Lee, 2006). Providing more information about gender, however, should also lead to heightened reliance on gender categories to predict behavior. In future research, it would be interesting to examine how young children integrate information about past experiences and social categories when both sources of information are highly salient. Using realistic examples of behavior, Berndt and Heller (1986) found that young children preferred to base inferences about the future on gender stereotypes rather than on past behavior; however, children in this

study were also given relatively limited information (see also Biernat, 1991). Thus, this remains an important area for future work.

These findings are also limited by the cultural context in which we conducted this research. Specifically, the age-related decline in the influence of gender-linked information documented in the present studies must be interpreted within the context of the present sample. All the data for these experiments were collected in a relatively liberal university town. Thus, it is possible that other cultural contexts support the maintenance of strict gender concepts throughout development. Interestingly, even in the relatively liberal cultural context sampled in the present work, we still found that young children hold relatively strong views regarding the influence of gender on behavior and reasoned about gender categories in a comparable manner to how they reasoned about animal species. In less egalitarian contexts, older children and even adults may continue to rely on gender to guide inferences about individual consistency when limited evidence is available. We interpret the present findings as indicating that whether individuals rely on a category-based strategy when making social inferences depends on the causal power attributed to the category for determining a wide range of behaviors. Whether gender is viewed in this manner should depend on age (as documented in the present work) as well as context (e.g., Crouter, Whiteman, & McHale, 2007).

In future research, it would also be interesting to examine how individuals reason about other social categories. We demonstrated that preschoolers appeal to membership in gender categories, but not more trivial differences based on shirt color, to guide their inferences. Based on our framework, we would expect that other social categories that individuals perceive as causal motivators of behavior to be used in a similar manner. For example, prior work suggests that young children attribute causal influence to categories based on race (Hirschfeld, 1996) and ethnicity (Astuti et al., 2004; Diesendruck & haLevi, 2006). Thus, we would expect children (and adults) to rely on these categories and others as well, to the extent that they view the categories as having a strong influence on a wide range of known and unknown behaviors.

In sum, these results inform two aspects of the development of social cognition, children's understanding of the mechanisms leading to individual consistency and the ways in which concepts of social categories influence children's expectations about the social world. These findings extend knowledge of how young children make sense of the social world, by pointing toward cues that preschoolers use when determining which behaviors are likely to be engaged in consistently over time. The present findings also document the importance of considering how young children's emerging theories of behavior and of social groups operate together to inform their explanations and predictions about the social world.

References

Aloise, P. A. (1993). Trait confirmation and disconfirmation: The development of attribution biases. *Journal of Experimental Child Psychology, 55*, 177–193.

Alvarez, J. M., Ruble, D. N., & Bolger, N. B. (2001). Trait understanding or evaluative reasoning? An analysis of children's behavioral predictions. *Child Development*, *72*, 1409–1425.

Astuti, R., Solomon, G. E. A., & Carey, S. (2004). Constraints on conceptual development. *Monographs of the Society for Research in Child Development*, 69, (2, Serial No. 277).

Berndt, T. J., & Heller, K. A. (1986). Gender stereotypes and social inferences: A developmental study. *Journal of Personality and Social Psychology, 50*, 889–898.

Biernat, M. (1991). Gender stereotypes and the relationship between masculinity and femininity: A developmental analysis. *Journal of Personality and Social Psychology, 61*, 351–365.

Boseovski, J. J., & Lee, K. (2006). Children's use of frequency information for trait categorization and behavioral prediction. *Developmental Psychology, 42*, 500–513.

Cain, K. M., Heyman, G. D., & Walker, M. E. (1997). Preschoolers' ability to make dispositional predictions within and across domains. *Social Development*, *61*, 53–75.

Choi, I., Nisbett, R. E., & Norenzayan, A. (1999). Causal attribution across cultures: Variation and universality. *Psychological Bulletin*, *125*, 47–63.

Crouter, A. C., Whiteman, S. D., & McHale, S. M. (2007). Development of gender attitude traditionality across middle childhood and adolescence. *Child Development*, *78*, 911–926.

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

Dozier, M. (1991). Functional measurement assessment of young children's ability to predict future behavior. *Child Development*, *62*, 1091–1099.

Ferguson, T. J., van Roozendaal, J., & Rule, B. G. (1986). Informational basis for children's impressions of others. *Developmental Psychology, 22*, 335–341.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday life*. New York: Oxford University Press.

Gelman, S. A., & Coley, J. D. (1991). Language and categorization: The acquisition of natural kind terms. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 146–196). New York: Cambridge University Press.

Gelman, S. A., Collman, P., & Maccoby, E. E. (1986). Inferring properties from categories versus inferring

categories from properties: The case of gender. *Child Development*, *57*, 396–404.

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.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*, 183–209.

Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the nonobvious. *Cognition*, *23*, 183–209.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*, 21–38.

Gnepp, J., & Chilamkurti, C. (1988). Children's use of personality attributions to predict other people's emotional and behavioral reactions. *Child Development*, *59*, 743–754.

Heyman, G. D., & Gelman, S. A. (1999). The use of trait labels in making psychological inferences. *Child Development*, *70*, 604–619.

Heyman, G. D., & Gelman, S. A. (2000). Preschool children's use of trait labels to make inductive inferences. *Journal of Experimental Child Psychology*, *77*, 1–19.

Hirschfeld, L. A. (1996). *Race in the making: Cognition, culture, and the child's construction of human kinds*. Cambridge, MA: MIT Press.

Kalish, C. W. (2002). Children's predictions of consistency in people's actions. *Cognition*, *84*, 237–265.

Kalish, C. W., & Shiverick, S. M. (2004). Children's reasoning about norms and traits as motives for behavior. *Cognitive Development*, *19*, 401–416.

Karniol, R., & Ross, M. (1976). The development of causal attributions in social perceptions. *Journal of Personality and Social Psychology*, *34*, 455–464.

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*, 233–252.

Levy, G. D., Taylor, M. G., & Gelman, S. A. (1995). Traditional and evaluative aspects of flexibility in gender roles, social conventions, moral rules, and physical laws. *Child Development*, *66*, 515–531.

Liu, D., Gelman, S. A., & Wellman, H. M. (2007). Components of young children's trait understanding: Behavior-to-trait inferences and trait-to-behavior predictions. *Child Development*, *78*, 1543–1558.

Maccoby, E. E. (1988). Gender as a social category. *Developmental Psychology*, *24*, 755–765.

Medin, S., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). New York: Cambridge University Press.

Miller, J. G. (1984). Culture and the development of everyday social explanation. *Journal of Personality and Social Psychology*, *46*, 961–978.

Miller, J. G. (1986). Early cross-cultural commonalities in social explanation. *Developmental Psychology*, *22*, 514–520.

Miller, P. H., & Aloise, P. A. (1989). Young children's understanding of the psychological causes of behavior: A review. *Child Development*, *60*, 257–285.

Morris, M. W., & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology*, *67*, 949–971.

Norenzayan, A., Choi, I., & Nisbett, R. E. (2002). Cultural similarities and differences in social inference: Evidence from behavioral predictions and lay theories of behavior. *Personality and Social Psychology Bulletin*, *28*, 109–120.

Rholes, W. S., & Ruble, D. N. (1984). Children's understanding of dispositional characteristics of others. *Child Development*, *55*, 550–560.

Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. New York: Mcgraw-Hill.

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

Solomon, G. A. (2002). Birth, kind, and naïve biology. *Developmental Science*, *5*, 213–218.

Sousa, P., Atran, S., & Medin, D. (2002). Essentialism and folkbiology: Evidence from Brazil. *Journal of Cognition and Culture*, *2*, 195–223.

Taylor, M. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development*, *67*, 1555–1571.

Taylor, M., Rhodes, M., & Gelman, S. A. (in press). Boys will be boys; cows will be cows: Children's essentialist reasoning about human and animal development. *Child Development*.

Uleman, J. S. (1987). Consciousness and control: The case of spontaneous trait inferences. *Personality and Social Psychology Bulletin*, *12*, 337–354.

Waxman, S., Medin, D., & Ross, N. (2007). Folkbiological reasoning from a cross-cultural developmental perspective: Early essentialist notions are shaped by cultural beliefs. *Developmental Psychology*, *43*, 294–308.

Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.

Wellman, H. M., & Miller, J. G. (2006). Developing conceptions of responsive intentional agents. *Journal of Cognition and Culture*, *6*, 27–55.

Yuill, N. (1993). Understanding of personality and dispositions. In M. Bennett (Ed.), *The development of social cognition: The child as psychologist* (pp. 87–110). New York: Guilford Press.

Yuill, N. (1997). English children as personality theorists: Accounts of the modifiability, development, and origins of traits. *Genetic, Social, and General Psychology Monographs*, *123*, 5–26.

Yuill, N., & Pearson, A. (1998). The development of bases for trait attribution: Children's understanding of traits as causal mechanisms based on desire. *Developmental Psychology*, *34*, 574–586.

Appendix A
Sample Items Study 1

Human preference	Here is Megan. A few days ago, Megan did two things—Megan went famming and Megan went lopping. Megan liked famming. Today, Megan is doing two things again. Again, Megan is going famming and lopping. What will Megan like today? Will Megan like famming, like last time, or will Megan like lopping?
Human choice	Here is Billy. A few days ago, Billy had a choice. Billy could either go rilling or Billy could go taffing. Billy went rilling. Today, Billy has a choice again. Again, Billy can either go rilling or taffing. What will Billy do today? Will Billy go rilling, like last time, or will Billy go taffing?
Animal preference	Here is Buttercup. A few days ago, Buttercup did two things. Buttercup went soting and Buttercup went tabing. Buttercup liked soting. Today, Buttercup is doing two things again. Again, Buttercup is going soting and tabing. What will Buttercup like today? Will Buttercup like soting, like last time, or will Buttercup like tabing?
Animal choice	Here is Torty. A few days ago Torty had a choice. Torty could either go vanning or Torty could go jerning. Torty went vanning. Today Torty has a choice again. Again, Torty can either go vanning or jerning. What will Torty do today? Will Torty go vanning, like last time, or will Torty go jerning?
Nonsense words	Pipping, famming, lexing, zazzing, lopping, gorping, tooping, fepping, jerning, flecking, daxing, vanning, spouding, soting, heesting, bemming, tabbing, sugging, pagging, sudding, shaining, flooming, molling, risting, fulfing, thaying, blicking, plizzing, rilling, hepping, taffing, pomming

Appendix B
Sample Items Study 2

Same-gender pair	Here are two boys. A few days ago, they did two things; they went bemming and they went spouding. This boy liked bemming. This boy liked spouding. Today, this boy is doing two things again. Again, this boy is going bemming and spouding. What will this boy like today? Will this boy like bemming, like last time, or will this boy like spouding?
Different-gender pair	Here are two children, a girl and a boy. A few days ago they did two things; they went shaining and they went risting. The girl liked shaining. The boy liked risting. Today, the girl is doing two things again. Again, the girl is going shaining and risting. What will the girl like today? Will the girl like shaining, like last time, or will the girl like risting?
Same-species pair	Here are two pigs. A few days ago, they did two things; they went flooming and they went pagging. This pig liked flooming. This pig liked pagging. Today, this pig is doing two things again. Again, this pig is going flooming and pagging. What will this pig like today? Will this pig like flooming, like last time, or will this pig like pagging?
Different-species pair	Here are two animals, a lion and a bear. A few days ago they did two things, they went thaying and they went plizzing. The lion liked thaying. The bear liked plizzing. Today, the lion is doing two things again. Again, the lion is going thaying and plizzing. What will the lion like today? Will the lion like thaying, like last time, or will the lion like plizzing?
Sample choice question (asked to adults only)	Here are two children, a girl and a boy. A few days ago they had a choice. They could either go shaining, or they could go risting. The girl went shaining. The boy went risting. Today, the girl has a choice again. Again, the girl can either go shaining or risting. What will the girl do today? Will the girl go shaining, like last time, or will the girl go risting?

Appendix C
Sample Items Study 3, Pilot Study

Human preference	Here is a boy. A few days ago, this boy had two things for a snack, a fep and a dax. This boy liked the fep. Today, this boy is having two things for a snack again. Again, this boy is having a fep and a dax. What will this boy like today? Will this boy like the fep, like last time, or will this boy like the dax?
Human fear	Here is a girl. A few days ago, this girl saw two animals, a toop and a fleck. This girl was afraid of the toop. Today, this girl sees two animals again. Again, this girl sees a toop and a fleck. What will this girl be afraid of today? Will this girl be afraid of the toop, like last time, or will this girl be afraid of the fleck?
Animal preference	Here is a monkey. A few days ago, this monkey climbed two kinds of trees, floom trees and pag trees. This monkey liked climbing floom trees. Today, this monkey is climbing trees again. Again, this monkey is climbing floom trees and pag trees. What will this monkey like climbing today? Will this monkey like floom trees, like last time, or will this monkey like pag trees?
Animal fear	Here is a mouse. A few days ago, this mouse saw two animals, a lex and a gorp. This mouse was afraid of the lex. Today, this mouse sees two animals again. Again, this mouse sees a lex and a gorp. What will this mouse be afraid of today? Will this mouse be afraid of the lex, like last time, or will this mouse be afraid of the gorp?

Appendix D
Sample Items Study 3, Main Study

Same-gender pair *(preference)*	Here are two boys. A few days ago, these boys played two games—the fulf game and the blick game. This boy liked the fulf game. This boy liked the blick game. Today, this boy is playing two games again. Again, this boy is playing the fulf game and the blick game. What will this boy like today? Will this boy like the fulf game, like last time, or will this boy like the blick game?
Different-gender pair *(fear)*	Here are two children, a girl and a boy. A few days ago, these children saw two animals, a toop and a fleck. The girl was afraid of the toop. The boy was afraid of the fleck. Today, the girl sees two animals again. Again, the girl sees a toop and a fleck. What will the girl be afraid of today? Will the girl be afraid of the toop, like last time, or will the girl be afraid of the fleck?
Same-species pair *(fear)*	Here are two turtles. A few days ago, they saw two animals, a pom and a hep. This turtle was afraid of the pom. This turtle was afraid of the hep. Today, this turtle sees two animals again. Again, this turtle sees a pom and a hep. What will this turtle be afraid of today? Will this turtle be afraid of the pom, like last time, or will this turtle be afraid of the hep?
Different-species pair *(preference)*	Here are two animals, a bear and a monkey. A few days ago, they climbed two kinds of trees, rist trees and shain trees. The bear liked climbing rist trees. The monkey liked climbing shain trees. Today, the bear is climbing trees again. Again, the bear is climbing rist trees and shain trees. What will the bear like climbing today? Will the bear like climbing rist trees, like last time, or will the bear like climbing shain trees?

Appendix E
Sample Items Study 4

Different-gender pair	See these two children? This one is a boy. This one is a girl. A few days ago, they did two things; they went bemming and they went spouding. The boy liked bemming. The girl liked spouding. Today, the boy is doing two things again. Again, the boy is going bemming and spouding. What will the boy like today? Bemming, like last time, or spouding?
Different-shirt-color pair	See these two children? This one is wearing a green shirt. This one is wearing a yellow shirt. A few days ago they did two things; they went shaining and they went risting. The one wearing a green shirt liked shaining. The one wearing a yellow shirt liked risting. Today, the one wearing a green shirt is doing two things again. Again, the one wearing a green shirt is going shaining and risting. What will the one wearing a green shirt like today? Shaining, like last time, or risting?