

Boys Will Be Boys; Cows Will Be Cows: Children's Essentialist Reasoning About Gender Categories and Animal Species

Marianne G. Taylor
Pacific Lutheran University

Marjorie Rhodes and Susan A. Gelman
University of Michigan

Two studies ($N = 456$) compared the development of concepts of animal species and human gender, using a switched-at-birth reasoning task. Younger children (5- and 6-year-olds) treated animal species and human gender as equivalent; they made similar levels of category-based inferences and endorsed similar explanations for development in these 2 domains. In contrast, 10-year-olds and adults treated gender and species concepts as distinct from one another. They viewed gender-linked behavioral properties as open to environmental influence and endorsed environment-based mechanisms to explain gender development. At all ages, children demonstrated differentiated reasoning about physical and behavioral properties, although this differentiation became more stable with age. The role of psychological essentialism in guiding conceptual development is discussed.

A number of theorists have compared children's and adults' reasoning about gender and other human social categories to their reasoning about animal species, proposing that people appeal to a notion of a category "essence" in their reasoning about both kinds of categories (Allport, 1954; Atran, 1990; Atran et al., 2001; Gil-White, 2001; Haslam, Rothschild, & Ernst, 2000; Hirschfeld, 1996; Rothbart & Taylor, 1992; Taylor, 1996). Psychological essentialism indicates that people treat certain categories as having an underlying reality or true nature that gives category members their identity and observable properties (Medin & Ortony, 1989). Evidence that young children, like adults, reason about categories as if they are determined by underlying

essences has led to the proposal that psychological essentialism results from early emerging cognitive biases that guide conceptual development (Gelman, 2003).

The proposal that children apply the same explanatory framework to understand both animal species and human social categories has fostered rich theoretical questioning and debate regarding the onset of psychological essentialism in children's biological and social reasoning. For example, do children transfer essentialist reasoning from their concepts of animal species to understand the social world (Atran, 1990; Gil-White, 2001), or is psychological essentialism instantiated separately in the biological and social domains (Hirschfeld, 1996)? How similar is children's reasoning about animal and human social categories with respect to their ideas about what might constitute a category essence and how an essence is obtained (Kanovsky, 2007)? Yet, despite the richness of these theoretical questions, little empirical work has directly compared children's reasoning about human social categories and animal species (for exceptions, see Astuti, Solomon, & Carey, 2004; Rhodes & Gelman, 2008). The goal of the present set of studies was to compare children's reasoning about animal and gender development regarding several aspects of psychological essentialism: the immutability of category identity, the inherent nature of

This research was supported by National Institute of Child Health and Human Development Grant HD-36043 to Gelman. We thank the parents, teachers, and children at the following institutions in Ann Arbor, Michigan: University of Michigan Children's Centers, the Jewish Community Center, St. Thomas Elementary School, and the YMCA. We also thank participating families and schools in Champaign-Urbana, Illinois, and surrounding areas: the University of Illinois Child Development Lab, the Learning Tree, Royal Elementary School, Schneider Elementary School, St. Matthew's Elementary School, St. John's School, Washington Elementary School, the Montessori School, the Don Moyers Boys and Girls Club, the University of Illinois Summer Science Program, the Thomas Paine Computer Camp, and the University of Illinois Summer Fitness Camp. We extend our appreciation to Cindy Andress, Mandy Davies, Galena Kline, Tracey Kramer, and Aisling O'Driscoll for their assistance with data collection and coding.

Correspondence concerning this article should be addressed to Marianne Taylor, Department of Psychology, Pacific Lutheran University, Tacoma, WA 98447-0003. Electronic mail may be sent to taylormg@plu.edu.

© 2009, Copyright the Author(s)
Journal Compilation © 2009, Society for Research in Child Development, Inc.
All rights reserved. 0009-3920/2009/8002-0012

category-linked properties, and the causal mechanisms responsible for development.

Category Immutability and Category Properties

We used a version of the “switched-at-birth” task, which has been commonly used to examine children’s beliefs about the immutability of category membership. In this task, children are told stories in which a baby (animal or human) is born a member of a certain category and then raised entirely by members of another category. For example, a baby pig is raised entirely by cows (Gelman & Wellman, 1991), or a human male infant is raised entirely by females (Taylor, 1996). The test questions ask what the baby will be like when it has grown up—will it develop the category-typical properties of its birth category or of the category members in its environment? Thus, these questions pit nature against nurture, asking children to reason about whether the category identity present at birth will lead to the development of typical category-linked properties (consistent with essentialist thought), or whether development will be influenced by the environment of upbringing.

In other studies, switched-at-birth tasks have also been used to tap children’s beliefs about biological inheritance and how category identity is acquired (Giménez & Harris, 2002; Hirschfeld, 1995; Johnson & Solomon, 1997; Solomon, Johnson, Zaitchik, & Carey, 1996; Sousa, Atran, & Medin, 2002; Springer, 1996; Waxman, Medin, & Ross, 2007). Because the present studies focus on gender (which differs from animal species and other social categories, such as race, in that it cannot be predicted based on characteristics of birth parents), we cannot speak directly to the important question of how children understand biological inheritance. Rather, in the present studies, children are directly told the category identity that characters have at birth, and we examine whether children predict that category membership will be stable over time and lead to the development of typical category-linked properties (Gelman & Wellman, 1991; Taylor, 1996). Thus, we assess the extent to which gender and animal species membership are treated as *inherent and stable* (Hall & Waxman, 1993; Macnamara, 1986), not the extent to which they are treated as *inherited*.

Prior work has demonstrated that children as young as 4 years old consistently predict that animals who are raised entirely by members of another species will retain their category identity across time and develop the physical and behav-

ioral properties typical of their birth category, not the category in their environment (Gelman & Wellman, 1991). Preschoolers do so even on more challenging tasks where they are not told the kind identity of the animals at birth, but have to determine category membership based on knowledge of the animal’s parents (Johnson & Solomon, 1997; Sousa et al., 2002; Waxman et al., 2007) or when they are asked to reason about unfamiliar properties (Sousa et al., 2002; Waxman et al., 2007). Essentialist reasoning about animal development has been found across diverse cultures in both children and adults, suggesting that representing animal species in an essentialist manner may be universal (Astuti et al., 2004; Atran, 1990; Atran et al., 2001; Sousa et al., 2002).

Whereas numerous studies have examined children’s beliefs about animal development using a switched-at-birth task, only one set of studies has examined beliefs about gender in this context. Taylor (1996) found that young children predicted that infants would develop the physical and behavioral characteristics associated with their birth gender category even if they were raised entirely by members of the other gender. Although preschoolers viewed certain properties as more resistant to environmental influence than others (e.g., physical vs. behavioral characteristics), they nonetheless allowed for little environmental influence overall. In these experiments, younger children attributed more predictive power to gender categories than did older children and adults, especially for behavioral properties; adults and 10-year-olds reliably predicted that characters would develop the behavioral properties associated with their environment of upbringing (i.e., the gender-atypical behavioral properties; see also Smith & Russell, 1984; Ullian, 1976).

Although only one study has used the switched-at-birth methodology to examine beliefs about gender, research using other experimental tasks provides converging evidence for children’s essentialist beliefs about gender. For example, preschoolers expect members of the same gender to share novel physical and behavioral properties, even when they have diverse appearances (Gelman, Collman, & Maccoby, 1986) or personalities (Diesendruck & haLevi, 2006). Young children also exaggerate gender group differences by denying or misremembering gender anomalies (Liben & Signorella, 1987), evaluating anomalies as morally wrong (Levy, Taylor, & Gelman, 1995), assuming that cross-gender roles are mutually exclusive (Deák & Maratsos, 1998; Florian, 1996), and emphasizing

within-group similarity in natural language (Gelman, Taylor, & Nguyen, 2004). Although adults display relatively more flexible beliefs about gender (Berndt & Heller, 1986; Biernat, 1991; Diesendruck & haLevi, 2006; Rhodes & Gelman, 2008; Taylor, 1996), elements of essentialist reasoning can be found in adults' concepts of gender as well (Haslam, Rothschild, & Ernst, 2002; Hoffman & Hurst, 1990; Prentice & Miller, 2006).

Based on this prior work, in the present studies we expected young children to treat human infants and baby animals in equivalent ways, as maintaining identity over time and developing the physical and behavioral characteristics associated with their birth categories. These predictions are consistent with prior work but before now have never been examined directly. By precisely comparing responses to questions about gender and animal development in within-subject (Study 1) and between-subjects (Study 2) designs, the present studies allow us to examine the extent to which gender is treated as equivalent to species kind. For example, do children view a preference for trucks as an inevitable consequence of being born a boy to the same degree as they believe that saying moo is an inevitable consequence of being born a cow?

We also examined developmental changes in how individuals reason about species and gender. Prior work indicates that older children and adults allow for more environmental influence on development, especially for behavioral (as compared to physical) properties. Yet the extent to which older children and adults allow environmental influence on behavioral properties for species and gender has not been compared. According to Atran (1990), children initially apply an essentialist framework to understand animal species and then later transfer this explanatory framework to social categories. From this perspective, we should find increasing convergence in how individuals represent gender and animal categories with age.

An alternate theoretical account of psychological essentialism suggests, however, that essentialism represents a more general mode of understanding that is instantiated separately in the biological and social domains (e.g., Gelman, 2003; Hirschfeld, 1996). From this perspective, psychological essentialism plays a particularly important role in early conceptual development, by guiding children to focus on the similarities across diverse instances of a category. With age, as children appreciate the importance of within-category variability, they may revise their concepts, and essentialist biases may play a less central role (Gelman & Coley, 1991).

We also go beyond prior work by examining the flexibility of children's beliefs. Particularly, we examine the extent to which children maintain their original predictions when presented with an alternate suggestion (e.g., Giménez & Harris, 2002). Inclusion of these questions allows for a stronger test of the prediction that children reason similarly about animal species and human gender by allowing more sensitive measurement of their beliefs.

Further, we extend prior work on gender concepts by examining the categories "male" and "female" separately. Some prior work suggests that young children may hold stricter gender stereotypes about boys than girls (e.g., Berndt & Heller, 1986; Biernat, 1991; Levy et al., 1995). For example, Levy et al. (1995) found that although children and adults reported that female behavior was flexible and open to cultural variation, they judged male gender role violations to be highly inflexible and just as unacceptable as moral transgressions. Differences in reasoning about male and female have not been examined previously in research on essentialism.

Children's Explanations

Another central aim of the present research was to compare the explanations that children provide for species- versus gender-linked development. In Study 1, we examined how children explained why animals and humans would develop certain properties. Of particular interest was whether and when children explicitly reference category identity. Further, we examined how children characterize category-property relations. Recent work in adult cognition has distinguished between two distinct ways of describing category-property relations: formal explanations, in which category identity provides an acceptable explanation for why an individual has a particular property (e.g., the fact that an animal is a dog can explain why it is four-legged), and association-based explanations, in which properties are viewed as strongly associated with a category but cannot be explained by reference to category identity itself (e.g., the fact that an animal is a dog cannot explain why it wears a collar, even though dogs generally wear collars; Prasada & Dillingham, 2006). Adults reliably distinguish between property-category relations that license formal explanations and those that license associational explanations, even when they judge the properties in question as equally probable in how often they apply to category members. Further, this distinction has implications for how adults reason about categories; for example, adults

have normative expectations regarding properties understood as formally linked to category membership (e.g., they believe that dogs *should* be four-legged; Prasada & Dillingham, 2006).

Drawing on this framework, we examined whether children's explanations for category-based predictions appear to be formal or associational, and whether these descriptions differ by property type, domain, and participant age. This analysis also speaks to the broader question of how children view the relations between categories and properties (Rakison & Oakes, 2003; Sousa et al., 2002). In Study 2, we examined children's explanations more explicitly by asking participants to judge a series of possible causal mechanisms.

Comparing children's explanations for gender-linked and species-linked development is central to identifying the extent to which they represent the social and biological domains in a similar manner. For example, if children have transferred a mode of reasoning developed specifically for understanding the biological world to reasoning about human social categories (Atran, 1990), we should expect to find a great deal of similarity in children's explanations for development in these domains. Alternately, if essentialism represents a more general mode of understanding with distinct instantiations in different domains (Gelman, 2003; Hirschfeld, 1995), we may find more variability in the explanations that children provide across domains. For example, we may find less appeal to birth when children explain gender than animal development (Gelman & Hirschfeld, 1999; Kanovsky, 2007). Although some prior work has examined how children justify their responses on switched-at-birth tasks, this has been done primarily on questions designed to tap children's understanding of inheritance (e.g., Johnson & Solomon, 1997), which is not directly relevant to the current work.

Overview

Studies 1 and 2 make use of switched-at-birth tasks to examine beliefs about animal kinds and human gender. In Study 1, we use a within-subject design to examine 5-year-olds', 10-year-olds', and adults' predictions about development, beliefs about property flexibility, and open-ended explanations for their choices. In Study 2, we use a between-subjects design to examine 6-year-olds', 8-year-olds', and 10-year-olds' predictions about development, beliefs about property flexibility, and endorsement of possible causal mechanisms. We selected these ages to span the range of ages

through which we expect to find developmental change in representations of gender (e.g., Taylor, 1996). This work is the first study of gender concepts to examine beliefs about flexibility on a switched-at-birth task, the first study of gender concepts to provide a systematic examination of children's explanations, and the first study of gender concepts to provide a head-to-head comparison with species concepts. Our use of converging methods (predictions, flexibility, and explanations) permits a particularly in-depth look at these issues.

Study 1

Method

Participants

Participants included 160 children of two age groups, recruited from preschools, elementary schools, and summer programs in two Midwestern university towns and rural surrounding areas. Younger children (referred to as "5-year-olds") were assigned to either an adoption condition or a control condition (see below) and ranged in age from 4;0 to 6;1 (adoption condition, $n = 68$, mean age = 4;10; control condition, $n = 32$, mean age = 4;11). The older children (referred to as "10-year-olds") were all assigned to the adoption condition, and ranged in age from 8;11 to 11;2 ($n = 64$, mean age = 9;11). Also, 32 college students participated in the adoption condition. There were equal numbers of females and males at each age and in each condition. Participants were primarily White, with some African American participants and a range of socioeconomic backgrounds.

Materials and Procedures

A female experimenter interviewed children individually in two 15- to 20-min sessions, approximately 1 week apart. Children answered questions about animals in one session (animal task) and humans in the other (gender task). Half of the boys and half of the girls at each age and in each condition completed the animal task first; half completed the gender task first. On the gender task, half of the boys and half of the girls at each age were asked about a female target, half about a male target (by random assignment).

Warm-up task. Younger children completed a short warm-up task during their first session. Children viewed a color photograph of a baby bear and learned that it had been raised from birth with a

different bear family. They were asked two questions about what the baby would be like when it grew up: "Does Tiltan eat honey or does Tiltan eat ketchup?" and "Is Tiltan brown or is Tiltan purple?" All children were able to answer these questions correctly.

Adoption condition. For the animal task, there were four separate stories presented in one of two orders. (The two orders—rabbit, cow, mouse, cat and cat, mouse, cow, rabbit—were balanced by age and sex of participants. A preliminary analysis revealed no significant order effects.) For each story in the animal task, children were shown a picture of a baby animal (e.g., a baby cow named Gana) and told that it had been raised from birth by animals of another species (e.g., a pig family, including a mother and one or more offspring; see Appendix A). The experimenter pointed to pictures of the newborn animal and the adoptive animal family during the story to maintain children's attention. The newborn pictures were color photographs of newborn animals with few characteristics of an adult member of the species. The pictures showed none of the features about which children were subsequently questioned. Children were asked several memory questions to ensure that they recalled and understood the story (e.g., "Is Gana a cow or a pig?" "Does Gana live with cows?" "Does Gana live with pigs?" "Are there any cows where Gana lives, besides Gana?"). If children responded incorrectly to any of these questions, the experimenter repeated the story before proceeding.

For the gender task, children were told one story about an infant who was raised from birth on an island with members of the other sex (see Appendix B). Children were told about a baby girl who went to live with her uncle on an island inhabited by boys and men or about a baby boy who went to live with his aunt on an island inhabited by girls and women. Children were shown color photographs of the infant, the other-sex relative, and the island. The same photo was used to represent the female and male infant. The infant had no distinguishing gender characteristics (e.g., the baby had very little hair, wore a diaper, and was labeled with an unfamiliar name: Zillah for the girl, Arza for the boy). The island picture showed an aerial view of Hawaii. A variety of small magazine clippings of children and adults were attached to the island photo and represented the island's inhabitants.

Prediction questions. Children were asked a series of questions about what properties the baby animal or human infant would have when it grew up (see Tables 1 and 2). There were four questions for each

of the four targets in the animal task (16 total) and 16 questions about the human child in the gender task, which were asked in a separate random order for each child. Questions were asked in a forced-choice format, such that children were asked to select between properties associated with the target's own species or gender category (category-based prediction) and properties associated with the environment of upbringing (environment-based prediction). For example, during the animal task, the experimenter asked, "When Gana has grown up, what does Gana say? Does Gana say 'moo' [category-based] or does Gana say 'oink' [environment-based]?" An example question from the gender condition is, "When Zillah is a big kid, what does Zillah like to do? Does Zillah like to sew [category-based] or does Zillah like to build things [environment-based response]?" The order of presentation for the two choices was randomized. The 16 items for both tasks included four different types of properties: behaviors, preferences, physical traits, and category traits (physical traits that explicitly mention the category label; these were designed such that the properties were identical [except for the category label] for the animal and gender tasks). For the gender task, children were asked four additional questions about the target's physical appearance (e.g., "When Zillah is a big kid, does Zillah have long hair or does Zillah have short hair?") for exploratory purposes. We excluded these questions from the analyses because there were no comparable questions for the animal task.

Category-based predictions were scored as 1; environment-based predictions were scored as 0. Preliminary analyses indicated few differences in the pattern of responses obtained for questions about preferences and behaviors, so answers to these questions were summed to form one score referred to as "predictions about behavioral properties." Similarly, there were few differences in children's predictions about category traits and physical traits, so these responses were summed to form one score referred to as "predictions about physical properties." Higher scores represent a greater number of birth category-based predictions.

Explanations. Open-ended questions were used to assess children's explanations for their predictions (adults were not asked to provide explanations). For each target, the experimenter asked children to provide explanations for one of each type of property, by asking "How come?" after the children responded. These items were randomly selected before the testing session. If children justified their answers for any of the other items, their responses

Table 1

Items for the Animal Task, Studies 1 and 2

Property type	Target	Environment	Properties
Behavioral	Rabbit	Monkeys	Eats carrots/bananas*
	Cow	Pigs	Says moo/oink
	Mouse	Dogs	Runs away from/chases cats
	Cat	Elephants	Licks its fur, sprays itself with water
	Rabbit	Monkeys	Likes to hop around/climb in trees
	Cow	Pigs	Likes to chew on grass/get dirty*
	Mouse	Dogs	Likes to eat cheese/chew on bones
	Cat	Elephant	Likes to drink milk/eat peanuts*
Physical ^a	Rabbit	Monkeys	Has long/short ears
	Cow	Pigs	Has a straight/curly tail
	Mouse	Dogs	Has round/floppy ears
	Cat	Elephant	Has whiskers/a trunk
	Rabbit	Monkeys	Has a rabbit/monkey body*
	Cow	Pigs	Has cow/pig blood inside*
	Mouse	Dogs	Has a mouse/dog brain
	Cat	Elephant	Has a cat/elephant heart*

Note. All items were used in Study 1; items marked with an asterisk were used in Study 2.

^aFour of these properties included explicit mention of the category label and were included in order to have one set of properties that was identical for the animals and humans (except for the label).

Table 2

Items for the Gender Task (Female/Male Target)

Property type	Properties
Behavioral	Plays with a tea set/toy truck*
	Plays dress-up/baseball cards*
	Will be a nursery school teacher/construction worker
	Collects dolls/tools and nails
	Likes to sew/build things
	Likes to put on make-up/go fishing*
	Wants to be a nurse/firefighter
	Wants to be a ballet dancer/football player
Physical	Can grow up to be a mommy/daddy
	Can grow up to be a wife/husband
	Will grow up to have breasts/a beard
	Voice will stay the same/get deep and low
	Has girl/boy blood inside*
	Has a girl's/boy's body*
	Has a girl's/boy's brain*
	Has a girl's/boy's heart

Note. All items were used in Study 1; items marked with an asterisk were used in Study 2.

were noted but are not included in the analyses. Two independent raters coded children's explanations from transcripts of the testing sessions. Responses were initially coded into four coding categories: category-based explanations, biological mechanisms, environment, and preferences (see Table 3). Subsequently, category-based explanations were further subdivided into three

subtypes: formal explanations, association explanations, and other (see Table 3). Initial interrater reliability was 96% and disagreements were resolved by discussion.

Flexibility. After recording the participant's initial response to the prediction questions, as well as the child's explanation (on selected questions), the experimenter next asked whether the target animal could also have the other property. For example, if on the initial prediction question the child said that the cow would say "moo," the experimenter asked, "When Gana has grown up, can Gana also say 'oink'?" Similarly, in the gender condition, if a participant said that a girl would like to sew when she grows up, the experimenter asked, "When Zillah grows up, would she also like to build things?" Children were asked to explain their answers for the same items that had been selected prior to the testing session. Items were scored as 1 if the participant said the character *could* have the other property and 0 if the participant said they could not. Items were summed (separately for physical and behavioral properties and for human and animal targets), such that higher scores indicate greater flexibility (i.e., more often indicating that other properties were possible).

Memory questions. To assess children's memory for story details, we also included questions concerning the target's environment, for example, "Who does Gana live with, cows or pigs?" These items were interspersed with the other items, with

Table 3
Coding Scheme for Open-Ended Explanations, Study 1

Explanation code	Definition and examples
Category	States that the target does/does not belong to a particular category or is a generic statement about the category. <i>Ex.</i> : "Because cows like to chew on grass, and he's a cow"
Formal	Subtype of category explanation that indicates that category membership explains the presence of a property. <i>Ex.</i> : "Because she's a girl"
Association	Subtype of category explanation that describes category-property associations. <i>Ex.</i> : "Because girls usually like dolls better"
Other category	Category-based explanations that did not fit criteria for <i>formal</i> or <i>association</i> were coded as "other." <i>Ex.</i> : "Because she's not a boy"
Biological mechanism	Refers to biological mechanism or biological consequences, even if incorrect. <i>Ex.</i> : "If it drank the pig mother's milk"; "Because he was not born a girl"
Environment	Refers to environmental influences, including learning or being taught, or environmental factor that prevents target from acquiring the property. <i>Ex.</i> : "Well, he's gonna be all around monkeys all his life, and his mom and his brothers and sisters and friends will be climbing in trees"
Preferences	Refers to the characters' preferences or desires. <i>Ex.</i> : "Because she loves to collect dolls"

three memory questions for the gender task and three memory questions for the animal task (one for each target).

Control Condition

The purpose of the control condition was to ensure that the properties used were ones which children found to be strongly associated with the target categories. The materials and procedures were identical to those for the adoption condition with the following exceptions. First, children were told that the babies were raised with members of the same species (animal task) or sex (gender task). For example, they were told that a baby monkey went to live with another monkey family, or that a baby girl went to live with her aunt on an island where there were only girls and women. Half of the children who completed the animal task heard stories about baby animals from the target species described in the adoption condition (rabbit, monkey, cow, pig); the remaining children heard stories about baby animals from the species associated with the environment of upbringing in the adoption condition (mouse, dog, cat, elephant). This procedure ensured that children were familiar with the properties associated with both the target animals and the animal families who adopted them (e.g., they heard about either a rabbit raised with other rabbits or a monkey raised with other monkeys). The corresponding questions for the two stories were identical (e.g., "Does X like to hop around or does X like to climb in trees?" "Does X have long ears or short ears?").

Results

Experimental Task Memory Questions

Children were highly accurate on the memory questions (92% correct). This indicates that the task was understandable to even the youngest participants.

Prediction Questions

The data from the prediction questions are presented in Table 4. We analyzed these data through a 2 (domain: animal, gender) \times 2 (property: physical, behavioral) \times 3 (age group: 5, 10, adult) \times 2 (sex of human target: boy, girl) \times 2 (participant sex: male, female) repeated-measures analysis of variance (ANOVA), with domain and property as within-subject variables. This analysis revealed a number of main effects and interactions, which were all subsumed under two significant four-way interactions: Domain \times Property \times Age \times Sex of Human Target, $F(2, 149) = 3.76$, $p < .05$, and Property \times Age \times Sex of Human Target \times Participant Sex, $F(2, 149) = 2.97$, $p = .05$. We examined these effects through a series of post hoc analyses. All post hoc contrasts that are discussed throughout this article were significant at $p < .05$ after Bonferroni corrections; we present Cohen's d as a measure of effect size for critical contrasts.

Comparing species and gender. For physical properties, there were no differences in children's or adults' predictions about species and gender. For behavioral properties, however, 5-year-olds,

Table 4

Mean Proportions of Category-Based Predictions and Flexibility Judgments by Age, Target, and Property Type, Study 1

	5-year-olds		10-year-olds		Adults	
	Physical	Behavior	Physical	Behavior	Physical	Behavior
Predictions						
Animal	.86** (.20)	.84** (.24)	.95** (.11)	.73** (.27)	.95** (.09)	.60 [†] (.32)
Male	.88** (.12)	.66** (.14)	.95** (.08)	.60* (.22)	.94** (.09)	.29** (.20)
Female	.86** (.18)	.52 (.23)	.92** (.14)	.37* (.34)	.90** (.18)	.18** (.18)
Flexibility						
Animal	.27** (.33)	.31** (.36)	.23** (.25)	.40 (.28)	.21** (.15)	.69** (.23)
Male	.19** (.32)	.38 [†] (.39)	.20** (.25)	.62* (.30)	.20** (.18)	.94** (.10)
Female	.29** (.35)	.68** (.36)	.16** (.17)	.68** (.35)	.32* (.31)	.97** (.07)

Note. Standard deviations are provided in parentheses. Higher numbers for predictions mean higher numbers of birth category-based predictions; higher numbers for flexibility judgments mean increased flexibility. Mean proportions were compared to the proportion expected by chance (.5) in a series of one-sample *t* tests.

[†]*p* < .10. **p* < .05. ***p* < .01.

10-year-olds, and adults all made fewer category-based predictions about the behavioral properties associated with gender (for both male and female targets) than the behavioral properties associated with animal species (*ds* = 0.87–1.40).

Comparing physical and behavioral properties. Five-year-olds did not make differentiated predictions about the physical and behavioral properties associated with animal species. In contrast, 10-year-olds and adults made fewer category-based predictions about behavioral than physical properties for animals (*ds* = 1.04, 3.86). On questions about gender, participants from all age groups made fewer category-based predictions about behavioral than physical properties, for both male and female targets (*ds* = 1.38–1.95).

Comparing across age. Predictions about physical properties did not change with age, for either human or animal targets. For behavioral properties, mean levels of category-based predictions declined with age, for both human and animal targets (see Table 4).

Sex of human target. Five-year-olds and 10-year-olds made fewer category-based predictions about the behavioral properties associated with the category “female” than the category “male” (age 5, *d* = 0.74; age 10, *d* = 0.81). There were no other effects of this variable.

Participant sex. There were two effects associated with participant sex. Five-year-old girls made more category-based predictions about the physical properties associated with “female” than 5-year-old boys did (*d* = 0.76). Also, 10-year-old girls made fewer category-based predictions about the behavioral properties associated with “female” than 10-year-old boys did (*d* = 0.99).

Individual Response Patterns

We next examined whether individual participants had consistent response strategies. Participants were designated as “category-based” if they made category-based predictions on at least six out of the eight questions for both physical and behavioral properties for each kind of target (with separate codes for questions about animals and humans). Alternately, participants were designated as “differentiated” if they made category-based predictions on at least six questions about physical properties and environment-based predictions on at least six questions about behavioral properties. Participants were designated as “environment-based” if they made environment-based predictions on at least six out of the eight questions for both physical and behavioral properties for each kind of target. However, only 2 participants (of 192) fell into this last category, so the environment-based pattern was not considered further. All other response patterns were coded as “other.”

Five-year-olds rarely showed a differentiated response pattern, more often showing category-based and “other” response patterns (animal task: 0 differentiated, 54 category-based, and 14 other; gender task: 4 differentiated, 28 category-based, and 36 other). These patterns shifted toward increasingly differentiated response patterns for 10-year-olds and adults, especially for gender (age 10, animal task: 5 differentiated, 41 category-based, and 18 other; age 10, gender task: 20 differentiated, 30 category-based, and 14 other; adults, animal task: 8 differentiated, 17 category-based, and 7 other; adults, gender task: 16 differentiated, 1

category-based, and 15 other). Chi-square analyses confirmed that the distributions of these codes differed by domain for each age group ($ps < .01$); category-based codes were more common in the animal than gender domain, whereas differentiated codes were more common in the gender domain. The distribution of codes also differed by age—for animals, $\chi^2(4) = 20.48$, $p < .001$, for gender, $\chi^2(4) = 40.16$, $p < .001$ —reflecting the increase in the number of differentiated codes with age.

Explanations

Next, we examined the explanations that children gave for their predictions. After each response was coded, we calculated the percent of total explanations attributed to each code, separately for environment-based and category-based predictions, by age, target, and property-type. Environment-based predictions were rarely provided on questions about physical properties (across less than 10% of trials), so these explanations were not examined.

Explanations for category-based predictions. For category-based predictions, participants often gave category-based explanations. This explanation code includes responses in which participants referenced category identity to explain their property predictions (see Table 3). No other explanation was given across more than 10% of each type of question; therefore, for questions on which participants made category-based predictions, we report data only on children's category-based explanations.

Total category-based explanations: We first examined the total number of category-based explanations for these predictions (including formal, associational, and other explanations). Five-year-olds gave category-based explanations on 76% of trials, and 10-year-olds gave category-based explanations on 77% of trials. Chi-square analyses revealed that the proportion of category-based explanations did not differ by age, property, or target (all $ps > .1$).

Formal and associational explanations: We next divided these category-based explanations into three subcategories. Formal explanations explicitly stated that the property was developed *because* of category membership, with no further qualifications (e.g., "because she is a girl"; "because it is a cow"). These were distinguished from statements of category–property associations (e.g., "because a lot of girls sew"; "because cows usually have straight tails"). Most explanations fit one of these two criteria (92%); the rest were coded as "other" (see Table 3). Table 5 presents the number of formal

Table 5

Subtypes of Category-Based Explanations for Category-Based Predictions, by Age, Domain, and Property Type, Study 1

	5-year-olds		10-year-olds	
	Formal	Associational	Formal	Associational
Animal physical	68	10	75	32
Animal behavior	49	35	37	34
Gender physical	58	9	69	20
Gender behavior	63	27	22	51

and associational explanations by age, domain, and property type. For 5-year-olds, the distribution of explanation type did not differ by domain (animal or gender) for either physical or behavioral properties ($ps > .5$). For 10-year-olds, the distribution of explanation types differed by domain for behavioral properties, $\chi^2(1) = 7.19$, $p < .01$; they gave a higher proportion of associational explanations for gender than for animal behavior. For both 5-year-olds and 10-year-olds, the distribution of formal and associational explanations differed by property type, for both animals—age 5: $\chi^2(1) = 16.78$, $p < .001$; age 10: $\chi^2(1) = 5.90$, $p < .05$ —and gender—age 5: $\chi^2(1) = 5.96$, $p < .05$; age 10: $\chi^2(1) = 36.61$, $p < .001$. As shown in Table 5, in each case children gave a higher proportion of formal explanations for physical than behavioral properties. There were two effects of age: (a) 5-year-olds gave a higher proportion of formal explanations for animal physical properties than 10-year-olds did, $\chi^2(1) = 7.51$, $p < .01$, and (b) 10-year-olds gave a higher proportion of associational explanations for gender behavioral properties than 5-year-olds did, $\chi^2(1) = 25.67$, $p < .001$.

Explanations for environment-based predictions. We next examined the explanations that children gave for their environment-based predictions. For 5-year-olds the most common code was for psychological causes, indicating that children or animals had environment-based properties because they wanted them or liked them (25% of questions). In contrast, for 10-year-olds, the most common code was for environmental mechanisms, indicating that children and animals had environment-based properties because they were taught to have them, learned to have them, wanted to fit in, or could not have been able to develop their category-based properties (e.g., a girl had to play football instead of do ballet because there would have been no ballet schools on the island; 89% of questions). Chi-squares confirmed that the explanations given by either 5- or 10-year-olds did not differ by target

($p > .5$), although the distribution of these explanations did differ by age, $\chi^2(6) = 141.77$, $p < .001$.

Flexibility

Next, we examined participants' responses to questions about property flexibility (Table 4) through a 2 (domain) \times 2 (property) \times 3 (age) \times 2 (sex of human target) \times 2 (participant sex) repeated measures ANOVA, with domain and property as within-subject variables. This analysis revealed a number of main effects, subsumed under interactions between domain and property, $F(1, 144) = 58.46$, $p < .001$, and property and age, $F(2, 144) = 49.03$, $p < .001$.

Comparing species and gender. For physical properties, there were no differences in participants' judgments about property flexibility for species and gender. For behavioral properties, however, participants viewed those associated with gender as more flexible than those associated with species (age 5, $d = 0.54$; age 10, $d = 0.69$; adults $d = 1.46$).

Comparing physical and behavioral properties. For both species and gender, participants judged behavioral properties to be more flexible than physical properties (age 5: animals, $d = 0.11$, gender, $d = 0.76$; age 10: animals, $d = 0.81$, gender, $d = 1.72$; adults: animals, $d = 2.54$; gender, $d = 3.81$).

Comparing across age. For physical properties, there were no changes in participants' judgments of flexibility related to age. For behavioral properties, however, flexibility judgments increased with age, for both species and gender (see Table 4). The interaction between age and property also indicated that, though participants of every age group predicted more flexibility for behavioral than physical properties, these judgments became increasingly differentiated with age.

The interaction among domain, property, and age was not significant. However, to ensure that participants of every age group made differentiated predictions about physical and behavioral properties for both species and gender, we conducted follow-up paired t tests separately by age and domain. These analyses confirmed that participants in all age groups predicted more flexibility for behavioral than physical properties for both species and gender ($ps < .05$). Also, in every age group, participants made similar predictions about flexibility for physical properties for species and gender, but predicted more flexibility for the behavioral properties associated with gender than species ($ps < .001$).

Control Group Analyses

The purpose of the control group was to ensure that the youngest participants associated the experimental properties with the relevant categories for both animals and human gender. On this task, 5-year-olds were asked to predict what properties animals or humans would have if they were raised in environments consisting only of members of their own birth category (species or gender). As expected, children readily made category-based predictions on these questions (95% of questions about animal physical properties, 92% of questions about animal behavioral properties, 91% of questions about gendered physical properties, and 82% of questions about gendered behavioral properties). A series of paired t tests indicated that children made fewer category-based predictions about the behavioral properties of gender than any other type of question ($ps < .01$). Children's responses on these questions were not influenced by whether the described character was male or female (for the gender task) or by participant sex.

The findings from the control group raised the possibility that younger children's differentiated predictions about gender-linked behavioral and physical properties (on the experimental tasks) reflected only baseline differences in their beliefs about the associations between the properties we selected and gender. To address this possibility, we selected the four gender behavior items (out of eight) about which children in the control condition made the strongest category-based predictions—89%, which does not differ from the proportion of times they made category predictions for gender-linked physical properties or either type of animal property ($ps > .1$). Next, we reran our analyses of children's property predictions and responses to flexibility questions in the experimental switched-at-birth task using this selected subset of items. We reasoned that if children in the experimental task had predicted more environmental influence for behavioral properties *only* because they viewed some of the behavioral properties as less strongly associated with gender categories to begin with, then young children should no longer differentiate between gender-linked physical and behavioral properties on this follow-up analysis. However, we found an identical pattern of results. Specifically, for property predictions, this analysis replicated the four-way interactions among domain, property, age, and sex of child target, $F(1, 123) = 7.54$, $p < .001$, and property, age, sex of child target, and participant sex, $F(1, 123) = 3.98$, $p < .05$.

Post hoc analyses confirmed the same pattern of findings as reported above, including that young children made more category-based predictions for gender-linked physical than gender-linked behavioral properties. Using this restricted set of gender behavior items also did not change our pattern of findings for analyses of responses to flexibility items; using this subset, participants of every age group judged the behavioral properties associated with gender as more flexible than the gendered physical properties, as well as more flexible than animal behavior ($ps < .001$).

Discussion

Study 1 revealed many similarities in how young children reason about species and gender. Participants of all ages viewed physical development as strongly constrained by category memberships for both species and gender. The youngest participants, however, also tended to make category-based predictions for the behavioral properties associated with gender, indicating that they view gender categories as more constraining, and more similar to animal species, than adults do. A similar pattern was found in our analyses of flexibility questions. The youngest children also equally often explained species-linked and gender-linked development with formal explanations, whereas older children gave fewer formal explanations when explaining gender-linked as compared to species-linked behavioral development. Despite the similarities in young children's reasoning about species and gender, however, their responses were not identical across domains. Even 5-year-olds viewed the behavioral properties associated with gender as more open to the environment and more flexible than the behavioral properties associated with species, though these predictions became more differentiated with age. We also found that, though 5-year-olds did not make differentiated predictions about physical and behavioral properties regarding animals, they did so regarding gender, indicating that the development of differentiated beliefs about gender preceded differentiated beliefs about animals.

In considering children's responses to the flexibility questions, it is important to acknowledge that children may have interpreted these questions differently for behavioral and physical questions. For behavioral properties, an affirmative response may mean either that children believe that the story characters would display *both* properties or that the character would display the suggested alternate property *instead* of the participant's original choice.

In contrast, for physical properties, to the extent that children view the choices as mutually exclusive (e.g., long vs. short ears), children could only interpret the flexibility questions as asking whether the character could display the suggested property *instead* of their original choice. Thus, children's greater acceptance of alternate suggestions on the flexibility questions about behavioral properties may relate to these differences in interpretation and should be considered cautiously. However, to the extent that our analyses of flexibility items are consistent with our findings on our other tasks (as is the case with older children and adults), this task may be interpreted as providing evidence that participants view physical development more than behavioral development to be constrained by category identity.

The finding that young children allow more environmental influence on gender-linked behavior than animal-linked behavior may suggest that "innateness" plays a stronger role in animal than social concepts, consistent with the proposal that essentialism is instantiated separately, and takes on different meaning, in biological and social domains (Gelman, 2003; Hirschfeld, 1996; Kanovsky, 2007). Also, the finding that concepts of gender and species become more (as opposed to less) differentiated with age is consistent with this more general conceptualization of essentialism, and not with the proposal that children transfer an essentialist framework from the biological to the social domain.

However, our analyses of younger children's explanations did not reveal differences across domains, as 5-year-olds commonly used the formal mode of explanation to account for both species-linked and gender-linked development. Prasada and Dillingham (2006) point out, however, that the formal mode of explanation makes no commitments about the mechanisms involved in mediating category-property relations. For example, the formal mode applies equally well to artifacts (e.g., "that sign is red because it is a stop sign") as to animals (e.g., "that animal is four-legged because it is a dog"), even though the mechanism that mediates the link between being a stop sign and being red is different from the mechanism that mediates between being a dog and being four-legged. Thus, the formal mode may, but need not, incorporate an expectation of a causal essence. This raises the possibility that although both gender-linked and species-linked properties are viewed as formally linked to category membership, the mechanisms that children believe mediate these links differ by domain. We examine this possibility directly in

Study 2, by presenting children with five possible specific mechanisms that could be responsible for the development of category-linked properties.

Study 2

Method

Participants

Participants included 264 children, randomly assigned to one of three conditions (animal target, female target, and male target), with the constraint that each condition had approximately equal numbers of males and females, from three age groups: 75 six-year-olds ($M = 6.3$, range = 4,10–6,10, $Ns = 21$ –28 per condition), 89 eight-year-olds ($M = 7.9$, range = 6,11–8,10, $Ns = 28$ –32 per condition), and 110 ten-year-olds ($M = 10.3$, range = 8,11–11,6, $Ns = 33$ –34 per condition). There were approximately equal numbers of boys and girls in the 8- and 10-year-old age groups. In the 5-year-old group, there were 16 girls and 10 boys in the animal target condition, 14 girls and 7 boys in the male target condition, and 15 girls and 13 boys in the female target condition. The sample was from the same population as in Study 1, but no child participated in both studies.

Materials and Procedures

Children in the animal target condition heard three different animal stories, presented in a separate random order for each child. Children in the female target condition heard a story about a baby girl; those in the male target condition heard a story about a baby boy. The stories and pictures were identical to those used in Study 1 (see Appendix A and Appendix B).

To reduce task length, we used only three of the animal targets from Study 1: the rabbit raised with monkeys, the cow raised with pigs, and the cat raised with elephants. We also included fewer property pairs: three behaviors and three physical traits per condition (see Tables 1 and 2). The physical traits we chose were the category traits from Study 1. As there were no significant differences between the physical traits and category traits in Study 1, we selected the category traits, so we would be able to compare them directly across the animal and human target conditions.

For each of the three animal targets, there were two questions, one concerning a behavior and one concerning a physical trait. For the gender task,

participants heard a single story and were asked three questions each for physical and behavioral properties. The order of presentation was balanced within subjects so that participants received each type of item first half of the time. For the female- and male-target conditions, items were presented in a separate random order for each child. Prediction, flexibility, and memory questions were asked and scored using the same procedures as in Study 1.

Causal Mechanisms

After making their initial predictions, children were asked to consider five different causal mechanisms and to determine whether each mechanism could explain how the target may have acquired the property: birth, wanting, environmental pressure, learning, and being taught. For example, if a child in the animal condition said that a baby rabbit raised with monkeys would grow up to eat carrots, the experimenter asked, “Is it because Javan was born that way?” (birth), “Is it because Javan really wanted to eat carrots?” (wanting), “Is it because the monkeys really wanted Javan to eat carrots?” (environmental pressure), “Is it because Javan learned to eat carrots?” (learning), and “Is it because the monkeys taught Javan to eat carrots?” (teaching). Children could answer affirmatively to more than one causal mechanism. These mechanisms were presented in a separate random order for each question for each child.

Results

Memory Questions

Children were highly accurate on the memory questions (99% correct), indicating that participants of all ages were able to follow the stories.

Prediction Questions

Table 6 presents the results. To examine the influence of domain and age on children’s predictions, we conducted a 3 (domain: animal, boy, girl) $\times 2$ (property: physical, behavioral) $\times 3$ (age: 6, 8, 10) $\times 2$ (participant sex: male, female) repeated measures ANOVA, with property type as a within-subject variable, followed by a series of post hoc analyses. This analysis revealed main effects of age and property, which were qualified by two 2-way interactions: Domain \times Age, $F(4, 246) = 3.63$, $p < .01$, and Property \times Age, $F(2, 246) = 15.90$, $p < .001$.

Table 6
Mean Proportions of Category-Based Predictions and Flexibility Judgments by Age, Target, and Property Type, Study 2

	6-year-olds		8-year-olds		10-year-olds	
	Physical	Behavior	Physical	Behavior	Physical	Behavior
Predictions						
Animal	.85** (.30)	.63 (.39)	.89** (.26)	.61 (.41)	.94** (.15)	.40 (.41)
Male	.92** (.18)	.79** (.34)	.85** (.29)	.61 [†] (.37)	.93** (.14)	.51 (.42)
Female	1 [†] (.00)	.77** (.29)	.85** (.28)	.55 (.44)	.79** (.31)	.22** (.38)
Flexibility						
Animal	.15** (.30)	.37 (.40)	.11** (.26)	.39 (.41)	.06** (.15)	.60 (.41)
Male	.08** (.18)	.21** (.18)	.15** (.29)	.38 [†] (.37)	.07** (.14)	.49 (.42)
Female	0 [†] (.00)	.23** (.29)	.15** (.28)	.45 (.44)	.21** (.31)	.78** (.38)

Note. Standard deviations are provided in parentheses. Higher numbers for predictions mean higher numbers of birth category-based predictions; higher numbers for flexibility judgments mean increased flexibility. Mean proportions were compared to the proportion expected by chance (.5) in a series of one-sample *t* tests.

*Significantly different from chance according to a binomial test, $p < .001$.

[†] $p < .10$. ** $p < .01$.

Comparing species and gender. The Domain \times Age interaction indicated that 6- and 8-year-olds made similar predictions in all conditions; that is, predictions did not vary by domain. In contrast, 10-year-olds made fewer category-based predictions about girls than about animals or boys ($d = 0.53$, $d = 0.70$).

Comparing physical and behavioral properties. The interaction between property and age indicated that children at each age group made more category-based predictions for physical than for behavioral properties (age 6, $d = 0.66$; age 8, $d = 0.79$; age 10, $d = 1.6$), and that category-based predictions about behavioral properties declined with age (see Table 6). The three-way interaction among age, target, and property was not significant, implying that the effect of property (i.e., more category-based predictions for physical than behavioral properties) held across age groups and conditions. To examine this effect more closely, we conducted paired *t*-tests comparing children's judgments about physical and behavioral properties separately for each age group and condition (domain). Consistent with the analysis above, in every age group and condition, participants made category-based predictions more often for physical than behavioral properties (all $ps < .001$).

Individual Response Patterns

We next examined individuals' overall response patterns. We classified individuals as category-based if they made category-based predictions on at least five out of the six total questions (summing behavioral and physical properties). We classified

individuals as differentiated if they showed the differentiated pattern (category-based predictions for the physical properties only) on at least five out of the six questions. Participants were classified as environment-based if they made environment-based predictions on five out of six questions. As in Study 1, however, very few children fit these criteria (1 six-year-old, 2 eight-year-olds, and 1 ten-year-old), so this pattern was not considered in analyses. All other response patterns were coded as "other."

Six-year-olds rarely showed a differentiated response pattern (4 total), more often showing category-based response patterns (36 total). These patterns shifted toward increasingly differentiated response patterns at ages 8 (12 differentiated, 32 category-based) and 10 (43 differentiated, 20 category-based). Chi-square analyses indicated that these distributions differed by age—for animal items, $\chi^2(2) = 10.18$, $p < .01$; for boy items, $\chi^2(2) = 10.19$, $p < .01$; for girl items, $\chi^2(2) = 22.28$, $p < .001$ —but did not differ by domain in any age group ($ps > .20$).

Explanations

Next, we examined children's explanations, separately for category- and environment-based predictions. For each type of prediction, we calculated the proportion of times children endorsed each explanation (see Table 7). The number of times children endorsed each explanation was divided by the number of times they made the prediction of interest. For example, a child who made category-based predictions on all three questions about physical

Table 7
Mean Proportions of Endorsements for Causal Explanations by Type of Prediction, Study 2

	6-year-olds	8-year-olds	10-year-olds
Category-based predictions about physical properties			
Born ^a	.96 (.18)	.93 (.26)	.97 (.16)
Want ^a	.50 (.41)	.20 (.33)	.04 (.13)
Learn ^a	.38 (.42)	.17 (.30)	.06 (.15)
Taught	.18 (.31)	.01 (.06)	.01 (.05)
Environment	.26 (.38)	.06 (.18)	0 (0)
Category-based predictions about behavioral properties			
Born ^a	.44 (.44)	.36 (.42)	.18 (.37)
Want ^a	.65 (.41)	.44 (.45)	.25 (.41)
Learn ^a	.53 (.43)	.38 (.42)	.14 (.32)
Taught	.17 (.28)	.03 (.13)	.03 (.17)
Environment	.20 (.32)	.06 (.19)	.04 (.17)
Environment-based predictions about behavioral properties			
Born	.02 (.07)	.01 (.07)	.07 (.36)
Want ^a	.08 (.24)	.05 (.16)	.22 (.36)
Learn ^a	.13 (.32)	.19 (.36)	.59 (.47)
Taught ^a	.08 (.26)	.11 (.27)	.32 (.42)
Environment ^a	.12 (.29)	.23 (.39)	.56 (.47)

^aIndicates explanation types that were endorsed often enough to be included in analyses.

properties and who indicated that targets had their category-based physical properties because they were born with them on two of these three questions was assigned a score of 2/3 (.66) for the “born” explanation. Because children could endorse more than one explanation for each of their predictions, it was possible for these scores to sum to more than 1. For each type of prediction, we first identified explanations that were endorsed relatively often; only explanation types that were endorsed more than 15% of the time across age groups for each type of prediction were included in subsequent analyses.

Explanations for category-based predictions. We conducted a 3 (domain: male, female, animal) \times 2 (property type: physical, behavioral) \times 3 (explanation: born, want, learn) \times 3 (age: 6, 8, 10) repeated measures ANOVA, with explanation type and property type as within-subject variables. This analysis revealed a number of main effects and interactions, which were all subsumed under two significant three-way interactions: Domain \times Property \times Explanation, $F(4, 264) = 4.09$, $p < .01$, and Property \times Explanation \times Age, $F(4, 264) = 7.95$, $p < .001$.

Comparing species and gender: For the most part, children endorsed similar mechanisms for human gender and species. The only domain effect was that children were more likely to endorse “want” to explain the behavioral properties of boys

than those of animals (boys, $M = .60$, $SD = .46$; animals, $M = .32$, $SD = .43$).

Comparing physical and behavioral properties: Children commonly said that characters had physical properties because they were born with them; children of each age group and in each condition (domain) endorsed “born” more often for physical than behavioral properties. These patterns were reversed for “wanting” and “learning”; children endorsed these explanations more often for behavioral than for physical properties (though for 6-year-olds’ judgments about “wanting,” this effect was marginal, $p < .1$).

Comparing across age: Endorsement of the “born” explanation for physical properties did not vary with age. In contrast, endorsement of “born” for behavioral properties, as well as endorsement of “want” and “learn” for both types of properties, declined with age. For behavioral properties, 6-year-olds endorsed each explanation type equally, but 8- and 10-year-olds endorsed “born” less often than “want” or “learn.”

Explanations for environment-based predictions. Because environment-based predictions for physical properties were made so rarely, we did not examine these explanations. We examined children’s explanations for their environment-based predictions for behavioral properties through a 3 (domain) \times 4 (explanation: want, learn, taught, environment) \times 3 (age: 6, 8, 10) ANOVA. We found significant effects for explanation, $F(3, 507) = 28.45$, $p < .001$, and age, $F(2, 169) = 20.53$, $p < .001$, as well as a significant Explanation \times Age interaction, $F(6, 507) = 5.77$, $p < .001$. Ten-year-olds were more likely to endorse each type of explanation than either 6- or 8-year-olds, indicating that young children were generally unable to explain their environment-based predictions, whereas older children readily appealed to environmental mechanisms. There were no main or interactive effects of domain, indicating that children endorsed similar mechanisms to explain animals and gender.

Flexibility

Next, we examined the influence of domain, property, and age on children’s beliefs about the flexibility of properties (see Table 6). We conducted a 3 (domain: animal, boy, girl) \times 2 (property: physical, behavioral) \times 3 (age: 6, 8, 10) \times 2 (participant sex: male, female) repeated measures ANOVA, with property as a within-subject variable. Children’s flexibility increased with age, $F(2, 245) = 12.12$, $p < .001$, and was greater for behavioral than

physical properties, $F(1, 245) = 167.21$, $p < .001$. These main effects were subsumed under two important interactions. The increase in flexibility associated with age was specific to behavioral properties, as evidenced by a significant Age \times Property interaction, $F(2, 245) = 15.84$, $p < .001$, and 10-year-olds viewed girls' properties as more flexible than the properties of boys or animals, as evidenced by a significant Age \times Domain interaction, $F(4, 245) = 3.64$, $p < .01$.

Discussion

In Study 2, only the oldest participants, 10-year-olds, reasoned differently about gender and animal species, and they did so only about the category "female," not "male." Younger children made similar levels of category-based predictions, gave similar judgments about flexibility, and endorsed the same causal mechanisms when reasoning about gender and species. In contrast, 10-year-olds viewed female gender-linked behavior as more open to environmental influence than animal-linked behavior.

Children's explanations suggested that they viewed physical properties as an inherent part of being born in a particular category, regardless of category type. In contrast, children viewed three explanations as possible mechanisms driving the development of category-based behavioral properties (birth, wanting, and learning), with no differences by domain. Finally, whereas younger children failed to endorse any mechanisms to explain their environment-based predictions, older children readily endorsed environmental explanations for these predictions.

General Discussion

We compared children's and adults' concepts of gender to their concepts of animal species, with a focus on the extent to which participants view these categories as constraining development and the mechanisms they endorse. The studies reveal four key findings: (a) gender concepts were initially treated by young children as equivalent to concepts of animal species in their innateness and inflexibility; (b) the gender concepts of "boy" and "girl" did not develop in tandem, but rather an understanding of environmental influences on "girls" emerged at a younger age than the comparable understanding for "boys"; (c) young children displayed surprising difficulty reasoning about environmental mechanisms that could influence species and gender-linked behaviors; and (d) nonetheless, even our

youngest participants had some differentiated beliefs about the development of physical and behavioral properties.

Comparing Concepts of Human Gender and Animal Species

Before age 10, children displayed similar concepts of species and gender, making a relatively high proportion of category-based predictions for both. This is a striking finding, as it means that young children treat the concepts of "boy" and "girl" as equivalent to species, in the extent to which features are inborn, inflexible, and intrinsically linked to category membership. For physical properties, children of all ages typically made category-based predictions for both domains. For behavioral properties, although younger children in Study 1 viewed gender-linked development as more open to the environment and more flexible than species-linked development, these differences were small and were not replicated in Study 2. Further, in both studies, young children applied a similar explanatory framework for understanding species and gender: The importance of category identities present at birth played a central role in their conceptual representations in both domains.

With age, however, children increasingly treated the two domains differently. Specifically, older children and adults made fewer category-based predictions about gender than species. Among 10-year-olds, this change was qualified by the gender of the target—they viewed female behavior as more open to environmental influence than male behavior—but, by adulthood, participants viewed both male and female behavior as more open to the environment and flexible than animal behavior. Further evidence comes from children's explanations. In Study 1, 10-year-olds (but not 5-year-olds) gave more formal explanations for explaining category-based animal behaviors than for explaining category-based gender behaviors. In fact, 10-year-olds provided associational explanations a majority of the time for gendered behaviors, indicating that they viewed gender-linked behaviors as associated with, but not explainable by, category identity.

In sum, our examination of the development of species and gender concepts across childhood suggests a threefold progression: (a) gender and species are initially treated as equivalent to one another; (b) next, "girls" are distinguished from boys and species; and (c) finally, gender concepts ("boys" and "girls") are consistently distinguished from species. The distinction between gender and species may in

part reflect children's beliefs about the behaviors in question, and not solely their concepts of the categories per se. For example, "likes to sew" (gender property) may be seen as a choice, whereas "runs away from cats" (species property) may be seen as an evolutionary imperative. This is further complicated by the possibility that how children interpret a behavior could interact with domain. For example, even when controlling for property content, older children may believe that behaviors are more physiologically motivated for animals but more volitionally motivated for people (e.g., rabbits like carrots because they are programmed to do so, whereas women like quiche because of socialization and choice). It would be interesting to vary the behaviors tested so as to examine this factor more directly (e.g., including physiological behaviors for gender, such as a boy having a deeper voice, analogous to a cow saying "moo").

The developmental pattern we obtained is inconsistent with the transfer model of conceptual development proposed by Atran (1990), which suggests that individuals transfer an essentialist framework from the biological to the social domain ontogenetically, and thus would predict increasing convergence over time. Instead, this developmental pattern is consistent with the proposal that essentialism represents more general cognitive biases, which become instantiated separately and follow distinct developmental trajectories in different domains (Gelman, 2003). However, it is also possible that transfer may happen earlier in development (i.e., below age 5 years, which was the youngest age tested in these studies). This could be tested by examining children's beliefs earlier in development.

Beliefs About Behavioral and Physical Development

Additional evidence that psychological essentialism is best characterized as a more general cognitive bias with a number of distinct instantiations comes from our analyses of children's reasoning about behavioral and physical properties. Children consistently distinguished physical and behavioral properties. In Study 1, 5-year-olds made more category-based predictions for gendered physical than gendered behavioral properties. They viewed behavioral properties as more flexible than physical properties for both species and gender. They also provided a higher proportion of formal explanations for physical than behavioral properties. In Study 2, 6-year-olds made more category-based predictions for physical properties than behavioral

properties and viewed behavioral properties as more flexible. They were also more likely to indicate that physical properties were caused directly by birth, whereas they endorsed three mechanisms—birth, wanting, and learning—to explain category-based behavioral development.

Although we found consistent differentiation between physical and behavioral properties among young children, these differences were typically small. In general, 5- and 6-year-olds viewed behavioral properties as predicted by birth categories, inflexible, and explainable by category membership, but did so less consistently than they did for physical properties. The most striking differences in children's reasoning about behavioral and physical properties were in their Study 2 explanations, where they endorsed "born" over 90% of the time for physical properties but only half the time for behavioral properties. Moreover, individual response patterns showed little differentiation among the youngest participants. Altogether, these results suggest that the differentiation revealed by the other tasks at the youngest age is more fragile and less consistent than that demonstrated by the older participants.

We will first discuss implications of these findings for understanding children's reasoning about physical properties, followed by implications for their reasoning about behavioral properties. Children's reasoning about physical development was consistent with psychological essentialism. They viewed physical development as constrained by category identities acquired at birth, consistent with the proposal that they view category identity as determined by an underlying, immutable, causal force. These findings do not speak directly to the biological nature of children's beliefs, as we did not examine the extent to which children understand birth as a biological process (see Gelman, 1990; Johnson & Solomon, 1997). Recent results from other studies, however, suggest that children expect that the essences that determine category identity are bodily in nature. For example, a number of cross-cultural studies have demonstrated that in communities where blood is a central focus of discourse, young children may come to believe that the category essence is, in fact, in the blood (Sousa et al., 2002; Waxman et al., 2007).

A related question is whether children's reasoning on these tasks is properly characterized as *causal*. Prasada and Dillingham (2006) note that formal explanations indicate only that category identity is responsible for a property, which may or may not

involve causal beliefs. For example, one might say that a sign is red because it is a stop sign, indicating that category identity is responsible for the property “red,” but it would not be correct to say that being a stop sign *caused* a given sign to be red. On the other hand, from an adult perspective it would be correct to say that being a dog *caused* an animal to have four legs (Prasada & Dillingham, 2006). Other recent studies of children’s reasoning on switched-at-birth tasks, however, suggest that children will use category identity to make predictions about novel properties on switched-at-birth tasks (when they cannot rely on prior associative knowledge), indicating that they attribute broad causal power to category essences (for a full argument of how switched-at-birth tasks provide evidence for causal essentialist reasoning see Sousa et al., 2002). On the other hand, Gelman and Bloom (2007) found that preschool children may represent properties as having an intrinsic connection to a category in the absence of understanding that the connection is causally mediated. Thus, the present findings suggest that children may reason about the physical development of animal species and gender in a causal essentialist manner, as well as that their expectations regarding the nature of the category essence in this context draw on their intuitive understanding of biology (see Gelman, 2003), although more direct work in these areas is clearly needed.

One possible interpretation of the findings regarding young children’s differentiation of physical and behavioral properties is that they view only physical, and not behavioral, development in essentialist terms. There are several features of these data, however, that we believe conflict with this interpretation. First, as noted above, young children reliably made category-based predictions about behavioral properties (though less often than about physical properties). Although one possibility is that they based their predictions about behavioral properties on their knowledge of category–property associations, as opposed to beliefs about the role of category identity in guiding behavioral development, this seems unlikely, given the high frequency of formal explanations that children provided in Study 1. Second, children typically endorsed intrinsic mechanisms—born, wanting, learning—to explain behavioral development in Study 2. Each of these mechanisms implicates the character as an active initiator of developmental change, rather than resulting from external forces. Thus, it appears that children believed that *wanting* or *learning* to engage in category-typical behaviors is inherent to

being a category member. In contrast, young children rarely endorsed externally motivated explanations—that the characters had been taught or that others wanted them to do these behaviors. Thus, children’s frequent category-based predictions about behavioral properties, as well as their explanations, support the interpretation that young children appealed to an essentialist framework for understanding behavioral development. Children’s more specific ideas about how essences mediate category–property relations, however, were somewhat different when they were asked to reason about behavioral, as compared to physical, development.

If young children view the development of behavioral properties as an intrinsically motivated part of category identity, why did they make fewer category-based predictions about these properties? There are at least two possible reasons. One possibility is that children believe that some behavioral properties are causally related to category membership, whereas others are only statistically associated. This would explain why young children provided a higher proportion of associational explanations for behavioral than physical properties in Study 1. Preliminary item-level analyses did not identify particular behavioral properties on which children were especially likely to make environment-based predictions, but individual children may have different beliefs about which behaviors are formally related to categories and which are not, resulting in an overall lower proportion of category-based predictions. In future work, however, it will be necessary to examine whether children do in fact make fewer category-based predictions for properties that they view as statistically, but not formally, connected to categories.

A second possibility is that young children believe that behavioral properties are formally connected to category membership but also incorporate knowledge of exceptions into their predictions. A number of authors have noted that children (and adults) may view a property as caused by a category essence, while at the same time recognizing that the property will not be manifest by all category members (Hollander, Gelman, & Star, 2002; Prasada, 2000). For example, individuals may believe that flying is an inherent property of birds while at the same time recognizing that not all birds fly. The reality that anomalies exist does not undermine the belief that, in general, birds fly because they are birds (Prasada, 2000).

Thus, essentialism does not imply that all members are alike (see also Gelman, 2003). Children are

also confronted with variation and anomalies (e.g., flightless birds; tomboys), which they seem to accept. Indeed, in natural language, generic noun phrases seem particularly well-suited to referring to variable categories, as they express generalizations that admit of exceptions (e.g., “birds fly” does not mean that all birds fly; Carlson & Pelletier, 1995). Within this framework, young children may believe that behavioral properties are causally related to categories, while at the same time acknowledging that not all individuals manifest such properties. Why they admit of exceptions more for some categories than others (e.g., “female” more than “male” or “animal”) remains an open question. It is possible that they have encountered more variation in feminine behaviors due, for example, to the more serious negative implications for males than for females who engage in counterstereotypical behaviors (Hort, 1989). It is also possible that the relatively higher status of males in our culture contributes to the greater essentializing of boys than girls (see Mahalingam, 2003, for analogous effects of status on essentialism regarding Indian caste).

Interestingly, environment-based predictions appeared to be inconsistent with younger children’s general explanatory framework for understanding species and gender development. Even though young children made environment-based predictions a substantial minority of the time, they failed to explain their choices in Study 1 and often rejected all of the offered explanations in Study 2. Thus, although they predict that *not all* category members have category-linked behavioral properties, they do not yet have a coherent explanation for why a category member will have an environment-linked behavioral property.

In contrast, older children appear less likely to appeal to an essentialist framework to explain behavioral development. Ten-year-olds readily generated and endorsed environment-based explanations (learning, teaching, and environmental pressure) to explain the development of behavioral properties. Thus, older children seem to view behaviors as linked to environmental, rather than wholly intrinsic, causes, especially when reasoning about gender.

Furthermore, older children were less likely than younger children to endorse explanations for their category-based predictions. For example, although they sometimes predicted that a boy raised with girls would retain boylike behavioral properties, they tended to reject all of the given explanations for why this might be. In this case, from our perspective, children have made a prediction that is

inconsistent with their theory; therefore, their working theory does not provide an explanatory framework for this question.

In sum, younger children appear to believe that both the behavioral and physical properties associated with category identity are inherent to membership, consistent with essentialist thought. The biggest areas of difference are in the mechanisms that they believe mediate these category–property relations. This diversity of mechanisms is consistent with a broader conceptualization of essentialism, whereby an essence is represented as an internal property, which may or may not be specifically biological and may or may not operate through specifically biological mechanisms (Gelman, 2003; Gelman & Hirschfeld, 1999; Waxman et al., 2007). More specific ideas about what constitutes an essence and how an essence operates are filled in throughout conceptual development, as are children’s beliefs about which categories (gender vs. species) and properties (physical vs. behavioral) are best understood from an essentialist perspective. The present and prior work suggests that children’s essentialist ideas about physical development are later incorporated into intuitive theories of biology, whereby birth and other bodily mechanisms (e.g., blood) take on special significance. In contrast, reasoning about behavioral properties appears to be less closely linked to biological mechanisms, and more open to revision with age.

Conclusions

The present findings suggest that younger children reason in an essentialist manner about both species and gender categories, but appeal to different specific mechanisms to explain the relation between category identity and different kinds of properties. Despite this diversity, younger children reference internal causes for both physical and behavioral properties and tend to reject environmental influence. In contrast, with age, children increasingly distinguish between gender and animal species and develop a new conceptual framework for understanding gender development that incorporates an important causal role for environmental factors. Future research should address the processes by which children update their naïve theories about development and the kinds of knowledge, experiences, and evidence that facilitate conceptual change. It will also be important to explore relationships among essentialist reasoning, gendered self-concepts, and behavior. Essentialist beliefs have the potential to influence children’s

developing gender identity as well as their attitudes and actions toward others.

References

- Allport, G. (1954). *The nature of prejudice*. Cambridge, MA: Addison-Wesley.
- Astuti, R., Solomon, G. E. A., & Carey, S. (2004). Constraints on conceptual development. *Monographs of the Society for Research in Child Development*, 69, 1–135.
- Atran, S. (1990). *Cognitive foundations of natural history: Towards an anthropology of science*. New York: Cambridge University Press.
- Atran, S., Medin, D., Lynch, E., Vapnarsky, V., Ucan Ek', E., & Sousa, P. (2001). Folkbiology doesn't come from folkpsychology: Evidence from Yukatek Maya in cross-cultural perspective. *Journal of Cognition and Culture*, 1, 3–42.
- 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.
- Carlson, G. N., & Pelletier, F. J. (Eds.). (1995). *The generic book*. Chicago: University of Chicago Press.
- Deák, G. O., & Maratsos, M. (1998). On having complex representations of things: Preschoolers use multiple words for objects and people. *Developmental Psychology*, 34, 224–240.
- Diesendruck, G., & haLevi, H. (2006). The role of language, appearance, and culture in children's social category-based induction. *Child Development*, 77, 539–553.
- Florian, J. E. (1996). Preschoolers' understanding of social role combinations. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 56, 7062.
- Gelman, R. (1990). First principles organize attention to and learning about relevant data: Number and the animate-inanimate distinction as examples. *Cognitive Science*, 14, 79–106.
- Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York: Oxford University Press.
- Gelman, S. A., & Bloom, P. (2007). Developmental changes in the understanding of generics. *Cognition*, 105, 166–183.
- Gelman, S. A., & Coley, J. D. (1991). Language and categorization: The acquisition of natural kind terms. In J. P. Byrnes & S. A. Gelman (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 146–196). Cambridge, UK: 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., & Hirschfeld, L. A. (1999). How biological is essentialism? In D. L. Medin & S. Atran (Eds.), *Folkbiology* (pp. 403–446). Cambridge, MA: MIT Press.
- Gelman, S. A., Taylor, M. G., & Nguyen, S. (2004). Mother-child conversations about gender: Understanding the acquisition of essentialist beliefs. *Monographs of the Society for Research in Child Development*, 69, 1–115.
- Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the nonobvious. *Cognition*, 23, 183–209.
- Gil-White, F. J. (2001). Are ethnic groups biological "species" to the human brain? *Current Anthropology*, 42, 515–554.
- Giménez, M., & Harris, P. L. (2002). Understanding constraints on inheritance: Evidence for biological thinking in early childhood. *British Journal of Developmental Psychology*, 20, 307–324.
- Hall, G. D., & Waxman, S. R. (1993). Assumptions about word meaning: Individuation and basic-level kinds. *Child Development*, 64, 1550–1570.
- 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*, 41, 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.
- Hoffman, C., & Hurst, N. (1990). Gender stereotypes: Perception or rationalization? *Journal of Personality and Social Psychology*, 58, 197–208.
- Hollander, M. A., Gelman, S. A., & Star, J. (2002). Children's interpretation of generic noun phrases. *Developmental Psychology*, 38, 883–894.
- Hort, B. E. (1989). Jane's gun and John's mascara: A difference in peer reactions to males and females who display cross-gender behaviors. *Dissertation Abstracts International*, 49, 4064.
- Johnson, S. C., & Solomon, G. E. A. (1997). Why dogs have puppies and cats have kittens: The role of birth in young children's understanding of biological origins. *Child Development*, 68, 404–419.
- Kanovsky, M. (2007). Essentialism and folksociology: Ethnicity again. *Journal of Cognition and Culture*, 7, 241–281.
- 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.
- Liben, L. S., & Signorella, M. L. (Eds.). (1987). *Children's gender schemata. New Directions for Child Development*, No. 38. San Francisco: Jossey-Bass.
- Macnamara, J. (1986). *A border dispute: The place of logic in psychology and philosophy*. Cambridge, MA: MIT Press.
- Mahalingam, R. (2003). Essentialism, culture, and power: Representations of social class. *Journal of Social Issues*, 59, 733–749.

- Medin, D., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). New York: Cambridge University Press.
- Prasada, S. (2000). Acquiring generic knowledge. *Trends in Cognitive Science*, 4, 66–72.
- Prasada, S., & Dillingham, E. M. (2006). Principled and statistical connections in common sense conception. *Cognition*, 99, 73–112.
- Prentice, D. A., & Miller, D. T. (2006). Essentializing differences between women and men. *Psychological Science*, 17, 129–135.
- Rakison, D. H., & Oakes, L. M. (Eds.). (2003). *Early category and concept development: Making sense of the blooming buzzing confusion*. New York: Oxford University Press.
- Rhodes, M., & Gelman, S. A. (2008). Categories influence predictions about individual consistency. *Child Development*, 79, 1270–1287.
- Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kinds? In G. R. Semin & F. Fiedler (Eds.), *Language, interaction, and social cognition* (pp. 11–36). Newbury Park, CA: Sage.
- Smith, J., & Russell, G. (1984). Development of gender constancy and selective attention to same-sex models. *Child Development*, 46, 849–856.
- Solomon, G. E. A., Johnson, S. C., Zaitchik, D., & Carey, S. (1996). Like father, like son: Children's understanding of how and why offspring resemble their parents. *Child Development*, 67, 151–171.
- Sousa, P., Atran, S., & Medin, D. (2002). Essentialism and folkbiology: Evidence from Brazil. *Journal of Cognition and Culture*, 2, 195–223.
- Springer, K. (1996). Young children's understanding of a biological basis for parent-offspring relations. *Child Development*, 67, 2841–2856.
- Taylor, M. G. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development*, 67, 1555–1571.
- Ullian, D. (1976). The development of conceptions of masculinity and femininity. In B. Lloyd & J. Archer (Eds.), *Exploring sex differences* (pp. 25–47). London: Academic Press.
- 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.

Appendix A

Sample Story for the Adoption Condition Animal Task (Study 1) and the Animal Condition (Study 2)

Now, I'm going to tell you about a baby cow named Gana. Look, here's a picture of Gana when

Gana was a baby. [Show Gana]. Right after Gana was born, when Gana was just a tiny baby cow, Gana went to live with a pig family. See, here are all the pigs in the pig family. [Show pig family]. Can you point to Gana? Can you point to all the pigs in the pig family?

The pigs loved Gana and took care of Gana. Gana lived with the pigs and became part of the pig family. Gana grew up with only pigs and had a happy life, but Gana never got to see another cow.

Control questions

Is Gana a cow or a pig? Does Gana live with cows? Does Gana live with pigs? Are there any cows where Gana lives besides Gana?

Test questions

Now, I have another picture to show you. This is a picture of Gana now that Gana has grown up. This picture is a surprise for later, but I want you to think about what Gana is like now that Gana has grown up, and I'm going to ask you some questions. (e.g., *prediction*: when Gana has grown up, what does Gana say? Does Gana say moo or does Gana say oink? *explanation*, Study 1, How come? *flexibility*: Can Gana also say X (opposite of child's response)?)

Appendix B

Story for the Adoption Condition Gender Task With a Female Target (Study 1) and the Female Target Condition (Study 2)

Now, I'm going to tell you about a baby girl named Zillah. Look, here's a picture of Zillah when Zillah was a baby. [Show Zillah]. Right after Zillah was born, when Zillah was just a tiny baby girl, Zillah went to live with her uncle on an island. See, here is Zillah's uncle. [Show uncle and island. Place uncle next to Zillah. Place island above Zillah and uncle]. Can you point to Zillah? Can you point to Zillah's uncle?

Now, I have to tell you something about this island. On this island there were only boys and men. Zillah was the only girl. Can you point to all the boys and men on the island? Zillah's uncle loved Zillah and took care of Zillah. Zillah lived with her uncle and became part of the uncle's family. Zillah grew up on the island with only boys and men and had a happy life, but Zillah never got to see another girl or woman.

Control questions

Is Zillah a girl or a boy? Does Zillah live with boys and men on the island? Does Zillah live with girls and women on the island? Are there any girls and women on the island besides Zillah?

Test questions

Now, I have another picture to show you. [Show closed cardboard door covering picture]. This is a picture of Zillah now that Zillah is a big kid. This

picture is a surprise for later, but I want you to think about what Zillah is like now that Zillah is a big kid, and I'm going to ask you some questions (e.g., *prediction*: When Zillah is a big kid, does Zillah play with a tea set or does Zillah play with a toy truck? *explanation, Study 1*: How come? *flexibility*: Can Zillah also play with a X (opposite of child's response)? How come?