Title
Exploring the origins and nature of social-group based inferences across early childhood

Author
Pronovost, Megan

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, MERCED

Exploring the origins and nature of social-group based inferences across early childhood

A dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Psychological Sciences

by

Megan Alyssa Pronovost

Committee in charge:

Professor Rose Scott, Chair
Professor Heather Bortfeld
Professor Jeffrey Gilger
Professor Anne Warlaumont

The Dissertation of Megan Alyssa Pronovost is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

Heather Bortfeld[1]

Jeffrey Gilger[1]

Anne Warlaumont[2]

Rose Scott, Chair[1]

University of California, Merced[1] & University of California, Los Angeles[2]
2019

Table of Contents

Acknowledgements

First, I would like to thank my advisor, Dr. Rose Scott, for being an incredible mentor to me throughout my graduate career. I am so thankful to have had you as my advisor. This dissertation simply would not have been possible without you.

I would also like to thank John Bunce and Erin Roby, who also supported me throughout my graduate career. I learned so much from the both of you, and my experience at UC Merced definitely would not have been the same without you.

I would especially like to thank my parents, who have instilled in me a passion for learning and hard work from a young age. Thank you for always encouraging me to pursue my goals, and never letting me give up on them.

Finally, I would like to thank Mike Pronovost. Thank you for listening to my many presentations, discussing my research with me, and dealing with me when I was stressed. Your love and support have been essential throughout this process, and I cannot thank you enough.

Curriculum Vitae

EDUCATION

July 2019 (expected)	PhD in Psychological Sciences (expected) University of California, Merced; GPA: 4.0 Advisor: Dr. Rose Scott
December 2015	M.A., Psychological Sciences, University of California, Merced
June 2012	B.S., Psychology, University of California, Davis; Highest Honors in Psychology; GPA: 3.81

PUBLICATIONS

Smith, M.A., & Scott, R. M. (under review). 20-month-old infants generalize preferences across members of a social group.

Smith, M. A., & Scott, R. M. (2017). 20-month-olds use social-group membership to make inductive inferences. *Proceedings of the 39th Annual Conference of the Cognitive Science Society*. London, UK: Cognitive Science Society.

Scott, R.M., #Roby, E., & #**Smith, M. A.** (2017). False-belief understanding in the first years of life. In J. Kiverstein (Ed.), *Routledge handbook of the Philosophy of the Social Mind* (pp. 152-171). New York, NY: Routledge. #Equal contributors.

Manuscripts in Preparation or Revision

Glenwright, M., Scott, R. M., & **Smith, M. A.**, & Hanlon-Dearman, A.C. (in revision). Children with autism spectrum disorder can attribute false beliefs in a spontaneous-response preferential-looking task.

CONFERENCE PRESENTATIONS

Smith, M.A. & Roby, E., & Scott, R. M. (June, 2018*). Toddlers' false-belief understanding is consistent across non-elicited-response paradigms with similar task demands.* Poster presented at the International Congress of Infant Studies, Philadelphia, PA.

Smith, M.A., & Scott, R. M. (May, 2018). *5-Year-olds' inductive inferences about ethnicity, gender, and religion: evidence from a diverse, rural sample.* Poster

presented at the Association for Psychological Sciences Annual Convention, San Francisco, CA.

Scott, R. M. & **Smith, M. A.** (April, 2017). *Hedging your (mistaken) bets: 23-month-olds can attribute false beliefs involving quantifiers.* Paper presented at the biennial meeting of the Society for Research in Child Development, Austin, TX.

Smith, M. A., & Scott, R. M. (April, 2017). *Exploring the origins of social-category-based inferences: Evidence from 37-month-olds.* Poster presented at the biennial meeting of the Society for Research in Child Development, Austin, TX.

Smith, M. A., & Scott, R. M. (April, 2017). *When do infants generalize dispositions across agents? Exploring the influence of social-group membership.* Paper presented at the biennial meeting of the Society for Research in Child Development, Austin, TX.

Smith, **M. A.**, & Scott, R. M. (May, 2016). *20-month-old infants can attribute prosocial and antisocial dispositions to agents.* Poster presented at the International Conference on Infant Studies, New Orleans, USA.

Smith, **M. A.**, & Scott, R. M. (March, 2015). *20-month-old infants expect members of a social group to share preferences*. Poster presented at the Biennial Meeting of the Society for Research in Child Development, Philadelphia, USA.

Smith, M. A., Lagattuta, K. & Sayfan, L. (2013). *The effects of past information on children's and adults' spontaneous production of generics about people and animals*. Poster presented at the Society for Research in Child Development, Seattle, USA.

INVITED TALKS

Spring 2014	Psychological Science Colloquium, "Do infants generalize preferences across members of a social group?" UC Merced

FELLOWSHIPS & AWARDS

Spring 2018	Graduate Dean's Dissertation Fellowship, $8250 University of California Merced
Spring 2018	Developmental Student Research Grant, $406 University of California Merced
Spring 2018	Jennifer Mendiola Memorial Travel Award, $175 University of California Merced

Spring 2017	Hatano Cognitive Development Fellowship, $1000 University of California Merced
Spring 2017	Developmental Student Research Grant, $1000 University of California Merced
Spring 2017	Student Travel Award, $300 Society for Research in Child Development
Spring 2017	Psychological Sciences Conference Travel Award, $350 University of California Merced
Spring 2016	William R. Shadish Award for Leadership and Service, $280 University of California Merced
Spring 2016	Developmental Psychology Research Fellowship, $500 University of California Merced
Spring 2015	Graduate Student Association Travel Award, $400 University of California Merced

TEACHING EXPERIENCE

Instructor of Record

Developmental Psychology	Summer 2018
Developmental Psychology	Summer 2017

Teaching Assistant

Cognitive Development	Spring 2016, Spring 2017, Fall 2018
Introduction to Psychology	Fall 2013, Spring 2015, Fall 2015, Fall 2017
Developmental Psychology	Spring 2014, Fall 2016, Spring 2018
Behavioral Genetics	Fall 2014

Guest Lecturer

Introduction to Psychology	"Prenatal Development, Infancy, and Childhood", Spring 2017

PROFESSIONAL EXPERIENCE

2013 – present	**Graduate Student Researcher**, Center for Early Cognition and Language, UC Merced. Supervisor: Rose Scott.
2012 – 2013	**Applied Behavior Analysis Therapist**, Central California Autism Center, Fresno, CA
2011 – 2012	**Undergraduate Research Assistant**, Mind-Emotion Development Lab, University of California at Davis.

LEADERSHIP AND SERVICE

Fall 2018	Reviewer, 2019 Biennial Meeting of the Society for Research in Child Development
Summer 2018	Student Volunteer, 30th Association for Psychological Sciences Annual Convention, San Francisco, CA.
Summer 2017	Student Volunteer, 39th Annual Conference of the Cognitive Science Society, London, UK
Fall 2016 – Spring 2017	Graduate Student Organizer, 4th Annual Conference for the Child and Family
Fall 2015 – Spring 2016	Graduate Student Organizer, 3rd Annual Conference for the Child and Family.
Spring 2016	Panelist, UC Merced Prospective Graduate Student Visit
Fall 2015	Graduate Student Organizer, Developmental Psychology Colloquium Series, UC Merced.
Spring 2015	Panelist, UC Merced Prospective Graduate Student Visit

CERTIFICATIONS AND SKILLS

November 2015	"Improving Teaching by Assessing Learning" Certificate, UC Merced Center for Engaged Teaching and Learning

September 2015	"Developing Teaching Strategies" Certificate, UC Merced Center for Engaged Teaching and Learning
January 2015	"Mastering the Classroom with 1st Generation College Students" Certificate, UC Merced Center for Engaged Teaching and Learning
January 2013	Applied Behavior Analysis Therapist, Level 1

PROFESSIONAL MEMBERSHIPS

2018 – present	Association for Psychological Sciences
2017 – present	Cognitive Science Society
2014 – present	International Congress of Infant Studies
2013 – present	Society for Research in Child Development
2014 – 2016	American Psychological Association

Abstract of the Dissertation

Categorization is a vital aspect of human cognition that helps guide learning and knowledge. However, when categorization is applied to social categories, it can have pernicious downstream effects such as stereotypes and prejudice. By preschool, children believe that members of a social category will share inherent, stable characteristics. Thus, it is important to understand when the tendency to use social categories to draw inferences about other people unfolds in early childhood. I began to address these issues in the current dissertation. Specifically, in three projects I examine the origins of social-group based inferences and how environmental influences shape these inferences across early childhood. Project 1 examined the types of characteristics that infants expect members of a social group to share. Twenty-month-old infants expected that a single individual would be consistent in her social dispositions. Infants did not generalize behavioral dispositions across members of a social group. However, additional results from Experiment 1 and 2 suggested that infants might have difficulty reasoning about social dispositions at this age. Project 2 showed that the manner in which parents discuss social groups influences how children learn about social categories and the beliefs that children form about social categories. Children that had generic statements read to them about a novel social category were more likely to view members of that category as being highly similar to one another than children that did not hear the generic statements. Project 3 demonstrated the relative salience of social categories in the environment, a more distal environmental influence, impacts the social categories that children attend to when making inductive inferences. Together, these studies shed light on the origins of social-group based inferences in infancy, and how the environment impacts these inferences across early childhood.

Chapter 1

Categorization is vital to human cognition. People view some categories as natural kinds that have deep, inherent natures that make members of the category similar to one another in both obvious and nonobvious ways. Known as psychological essentialism, viewing categories as natural kinds can lead people to generalize prior knowledge about a category to novel entities and situations and facilitate learning (Gelman & Markman, 1987; Rhodes & Mandalaywala, 2017). For example, infants expect perceptually dissimilar objects to share internal characteristics if the objects are labeled with the same noun (Graham, Kilbreath, & Welder, 2001). This can also be extended to biological categories such as animals. By identifying that a novel animal belongs to the familiar category "dog", one can infer that it will likely display common properties of that category (e.g., is alive, plays fetch, etc.).

This type of reasoning can also apply to kinds of people: Adults tend to assume that social categories capture fundamental, inherent similarities amongst collections of individuals and thus use prior knowledge about a social category to make inductive inferences about the physical, psychological, and behavioral properties of novel group members (e.g., Agerström, Björklund, Carlsson, & Rooth, 2012; Diesendruck & HaLevi, 2006; Leslie, Cimpian, Meyer, & Freeland, 2015; Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012; Rhodes & Mandalaywala, 2017; Storage, Horne, Cimpian, & Leslie, 2016). For example, by learning that a firefighter helps in emergencies, one can then infer that another firefighter will also display the same characteristics.

However, using this type of reasoning with social categories may cause people to use features such as surface similarity to infer intrinsic characteristics that are not necessarily true of the category (e.g., Haslam, Rothschild, & Ernst, 2002; Hirschfeld, 1996; Leslie, 2017; Prentice & Miller, 2007; Rhodes, Leslie, & Tworek, 2012; Rothbart & Taylor, 1992). This in turn can give rise to prejudice and stereotypes with profound negative consequences. For instance, Arab job applicants receive fewer job interview invitations and need to appear warmer and more competent to have the same odds at getting a job interview (Agerström et al., 2012), teachers expect girls to have lower cognitive achievement than boys (Tindall & Hamill, 2004), and teachers expect lower achievement from African-American youths than from Caucasians (Richman, Bovelsky, Kroovand, Vacca, & West, 1997). It is therefore critical to understand how the tendency to use social categories to draw inferences about other people emerges.

By preschool, the tendency to view social category members as sharing deep, inherent properties endowed with rich, inductive potential is well established (e.g., Bigler, Jones, & Lobliner, 1997; Diesendruck & HaLevi, 2006; Waxman, 2013). From a young age, children spontaneously categorize individuals as members of a social group and use these categories to guide their expectations about the world. Children use group membership to draw a variety of inferences such as others' preferences for objects and activities (e.g., Diesendruck & HaLevi, 2006; Gelman, Collman, & Maccoby, 1986; Martin, Eisenbud, & Rose, 1995), friendships (Shutts, Roben, & Spelke, 2013), as well as knowledge and beliefs (Birnbaum, Deeb, Segall, Ben-Eliyahu, & Diesendruck, 2010;

Taylor, 1996). For example, 5-year-old children expect that members of the same ethnicity will prefer the same activities (Diesendruck & HaLevi, 2006) and preschool-aged children expect that members of the same, but not the opposite, sex will prefer the same toys (Martin et al., 1995). This suggests that the tendency to view members of a social group as sharing inherent properties emerges at a young age. Because this tendency is robust by preschool, it is possible that children may use social group membership as a basis of their inferences even earlier in life. Thus, it is important to understand when the tendency to use social categories to draw inferences about other people unfolds in early childhood.

In my dissertation, I began to address these issues. Specifically, in three projects I examine the origins of social-group based inferences and how environmental influences shape these inferences across early childhood.

In Project 1 of my dissertation, I examined the nature of social-group based inferences in infancy. Considerable research suggests that the ability to detect social groups emerges in infancy, but less research has examined when infants assume that members of a social group will be similar to one another or begin to make inductive inferences based on social group membership. In a recent study, colleagues and I showed that by 20 months, infants expect members of the same social group to share food preferences (Smith & Scott, 2017). This suggests that the tendency to use social group membership to make inductive inferences emerges early. However, many questions remain about the nature of these inferences. In particular, older children and adults use group membership to make inductive inferences about a variety of characteristics such as object and activity preferences, personality traits, behaviors, friendships, and beliefs (Diesendruck & HaLevi, 2006; Martin et al.,1995; Shutts et al., 2013). In addition to food preferences, do infants expect members of a social group to share other characteristics?

In Project 1, I began to address this question by examining whether infants expect members of a social group to share social dispositions. By preschool, children expect a person's personality traits to remain stable over time (Boseovski & Lee, 2006). Similarly, kindergartners and adults expect that someone who previously performed positive behaviors would likely engage in positive behaviors in other situations (Heyman & Gelman, 1998). Infants are also attentive to behavioral dispositions such as prosocial and antisocial behaviors. For example, 6- to 10-month old infants use the valence of an agent's social actions to predict how others will react to that agent, and to guide their own interactions with that agent. Given this evidence that infants attend to social dispositions from early in life, it is possible that this may be a characteristic that infants attach to social group membership. Project 1 examined whether infants, like older children and adults, expect group members to share behavioral dispositions.

Although the research by Smith and Scott (2017) suggests that the tendency to make social-group based inferences emerges early, in infancy, it is also clear that the particular inferences that children make about social groups are learned. To return to my earlier example, obviously children are not born knowing that firefighters help people in emergencies: this is something that they must learn. Support for this comes from research examining children's understanding of a prevalent social category, race. Children are more likely to use race as a marker of group membership with increasing age, and children's tendency to view race as stable varies depending on the community the child is

exposed to and the child's own racial background (Kinzler & Dautel, 2012; Roberts & Gelman, 2015). Thus, it is clear that learning influences children's reasoning about social categories.

Project 2 explored how children may learn about social groups in their environment. It has long been assumed that the manner in which social groups are discussed in a child's environment can influence children's expectations about social categories (Bigler & Liben, 2007). For example, after 5- and 9-year-olds either heard the phrases "Rosie eats carrots whenever she can" or "Rosie is a carrot-eater," the children who heard the latter judged that Rosie would display the stable characteristic of liking to eat carrots in the future (Gelman & Heyman, 1999). Additional research has shown that particular types of statements such as generic statements may influence beliefs about a social category. Adults and 4-year-olds who heard generic statements describing a novel social category (e.g., "Zarpies like to eat flowers") were more likely to develop essentialist beliefs about the category than those who heard non-generic statements (e.g., "This one likes to eat flowers"; Rhodes et al., 2012).

However, much of this prior work investigating how children may learn through parental statements has been conducted with older children who already essentialize robustly. The one prior study that has looked at younger children was based on experimenters reading to children in a lab setting (Rhodes et al., 2012). Research has yet to examine whether parental input shapes the development of essentialist beliefs in younger children, and if so, the particular aspects of parental input that might contribute to the development of essentialist beliefs at early ages. Project 2 investigated whether a natural conversation between parents and children could influence children's beliefs about a social category. Parents and children read a picture book about a novel social group, and I examined whether and how parental input impacts children's subsequent inferences about that social group.

The results of Project 2 shed light on how proximal environmental factors (i.e. conversations with parents) affect children's social-group inferences. Project 3 explored the impact of a more distal environmental influence, the salience of particular social categories in a child's environment. The specific inferences that children make about social groups, and the attributes children attend to when making these inferences, vary with age and culture. For example, 4-year-olds in Israel pay more attention to ethnicity (i.e. Jewish or Arab) than gender when making inductive inferences about individuals (Birnbaum et al., 2010). In contrast, children in Northern Ireland tend to base their inferences on religion (i.e. Catholic or Protestant) over other social categories such as gender and ethnicity (Smyth, Feeney, Eidson, & Coley, 2017). These results suggest that aspects of the environment shape children's inferences.

However, the bulk of this research has been conducted in other countries such as Israel, where the salience of particular social groups is intertwined with violent, historical conflicts. Few studies have examined reasoning about ethnicity and religion in places where these groups are not involved in such high levels of conflict, such as in the United States. Studies in the United States have instead focused on how children use social categories based on gender and race. Research has yet to investigate how children in the United States use ethnicity information when reasoning about social categories. The few studies that have been conducted on religion have focused on primarily middle- to upper-

class White samples (Rhodes & Gelman, 2009; Shutts et al., 2013; Smyth et al., 2017). Therefore, the nature of social-group based inferences in children from a more culturally, ethnically, and economically diverse sample is unknown.

Project 3 addressed this issue by replicating a well-known study that was conducted in Israel (Diesendruck & HaLevi, 2006). The relative inductive potential of social categories (i.e., ethnicity, gender, religion) was investigated with 5- to 7-year-olds from a culturally, ethnically, and economically diverse area. Based on previous research, we expected that children would use gender over other social categories as a basis for their inductive inferences (e.g., Gelman et al., 1986; Rhodes & Gelman, 2009; Taylor, 1996; Taylor, Rhodes, & Gelman, 2009). It was unclear how children would use religion and ethnicity information but we hypothesized that children in Merced County would be less likely to make inferences based on ethnicity or religion because these categories are less salient in their environment.

Together, these studies shed light on the origins of social-group based inferences in infancy, and how the environment shapes these inferences across early childhood. Understanding the contexts that promote social-category generalization will aid in implementing measures to reduce the negative impact of stereotypes and prejudice.

Chapter 2

Categorization is vital to human cognition. Forming categories allows us to organize information efficiently and quickly, guides our expectations when encountering category members, and facilitates learning (Gelman, 1988; Medin, Ojalehto, Waxman, & Bang, 2015). For example, upon identifying that a novel entity belongs to a familiar category (e.g., *bear*), one can infer that it likely possesses common properties of that category (e.g., it is alive, hibernates, etc.). Categories not only apply to biological kinds, but to kinds of people as well. Adults tend to assume social categories (e.g., doctors, women) capture fundamental, inherent similarities amongst collections of individuals and thus use prior knowledge about a social category to make inductive inferences about the physical, psychological, and behavioral properties of novel group members (e.g., Agerström et al., 2012). However, less is known about whether infants use social-group membership to make predictions about the characteristics of others.

In early infancy, children begin to notice features that are correlated with group membership and use them to recognize whether others are similar or dissimilar to themselves. For example, infants as young as 3 months prefer same-race faces (Kelly et al., 2005), and 10-month-olds are more likely to accept toys from individuals who speak their own language (Kinzler, Dupoux, & Spelke, 2007). Infants are also more likely to accept foods previously endorsed by a speaker of their native language (Shutts, Kinzler, McKee, & Spelke, 2009), and are more likely to imitate actions produced by a native-language speaker (Buttelmann, Zmyj, Daum, & Carpenter, 2013; Howard, Henderson, Carrazza, & Woodward, 2015). In addition to facial and linguistic features, infants attend to a variety of other features that could indicate an individual is a member of their own group. For example, infants attend to features such as food preferences and clothing when reasoning about group membership (Mahajan & Wynn, 2012). Together, these results suggest that infants notice attributes that mark group membership and similarities to the self.

Not only do infants detect attributes correlated with group membership, recent evidence suggests that infants also use social-group membership to draw inferences about novel group members. Research by Smith and Scott (2017) has shown that 20-month-old infants expect members of a social group to share food preferences. Infants were first introduced to two novel social groups marked with costumes and labels. Then, infants saw a member of one of the social groups demonstrate a preference for one of two foods. In the test trial, infants saw a member of either the same social group or a member of a different social group pick either the same food or a different food. Infants expected that members of the same, but not different, social groups would pick the same foods. This suggests that the tendency to use social-group membership to make inductive inferences emerges early.

However, considerable questions remain about the nature of infants' inferences. Previous research suggests that older children use group membership to make inductive inferences about a variety of characteristics such as preferences for objects and activities, personality traits, behaviors, friendships, and beliefs (Boseovski & Lee, 2006; Diesendruck & HaLevi, 2006; Martin et al., 1995; Shutts et al., 2013). For instance, kindergartners expect that members of the same social category (e.g., individuals who are the same ethnicity) will share personality traits such as becoming angry at similar things,

or being afraid of the same things (Birnbaum et al., 2010). Do infants also expect members of a social group to share characteristics besides food preferences?

One thing that children are sensitive to from early in life is social dispositions. Infants use the valence of an agent's social actions to predict how others will react to, and guide their own interactions with, that agent (Hamlin, Wynn, Bloom, & Mahajan, 2011). This suggests that infants are attentive to prosocial and antisocial behaviors and the distinction between them early in development. This raises the possibility that infants may make group-based inferences about behavioral dispositions similarly to characteristics like food preferences.

However, such inferences require some understanding of prosocial and antisocial behaviors as stable characteristics, and the extent to which infants see prosocial/antisocial behaviors as a stable property of an agent remains unclear. The fact that infants will approach someone who was previously nice hints that infants might expect her to be nice again (Hamlin et al., 2011). Yet, direct tests of this have found mixed results (e.g., Duh, Wang, & Goldman, 2017; Koch, Kenward, Fawcett, & Gredebäck, 2015). For instance, 17-month-old infants who saw an agent steal another agent's toy anticipated she would steal again (Koch et al., 2015). But infants who saw the agent give a toy also anticipated she would later steal, making it unclear whether infants truly expected consistent behavior (Koch et al., 2015). Do infants see behavioral dispositions (e.g., being prosocial/antisocial) as a stable property of an individual? If so, do infants generalize behavioral dispositions across individuals in the same social group?

To address these questions, twenty-month-old infants were tested in a violation-of-expectation task using a procedure adapted from Smith and Scott (2017). Infants saw one member of a social group behave either prosocially or antisocially three times. In the test trial, they either saw the same agent as in the familiarization trial, or a different agent from that same social group. This agent behaved prosocially once and antisocially once. If infants expect a single agent to be consistent in her behavioral disposition, then they will expect the individual to show the same disposition in the test trial as in the familiarization trial, and will find it unexpected and look longer when she behaves differently (e.g., if she is prosocial in familiarization and antisocial in test) instead. Similarly, if infants expect members of a social group to share behavioral dispositions, then they will expect members of the same social group to behave similarly, and will find it unexpected and look longer when they behave differently (e.g., one member of the social group is antisocial and the other is prosocial) instead.

Experiment 1

Method

Participants. 30 healthy infants participated (14 female, 16 male; M_{age}= 20 months, 5 days, range 18 months, 9 days – 22 months, 5 days). Another 19 infants were tested but excluded because they failed to complete the experiment (15), because they did not watch the initial phase of one or both of the test trials (3), or because the child had a developmental disorder (1). Approximately equal numbers of infants were randomly assigned to each combination of test agent (same-agent, different-agent) and familiarization valence (prosocial, antisocial).

The infants' names in this and the following experiment were obtained from birth records provided by the California Department of Public Health, as well as from a

database of parents who had previously expressed interest in participating in research studies with their children. Parents were offered reimbursement for their transportation expenses, and their infant was given a small gift (book or t-shirt) for participating. Parents provided written informed consent for their infant's participation.

The racial and ethnic composition of the final sample for both experiments was 76% Caucasian, 2% Asian, 2% American Indian or Alaska Native, 4% Black or African-American; an additional 2% chose 'other race', 5% selected more than one race, and 9% chose not to respond. 33% of the sample identified as Hispanic or Latino, 65% identified as not Hispanic or Latino, and 2% chose not respond. In lieu of income information, we recorded the highest level of education reported by either parent: 2% completed less than high school, 18% completed high school, 20% completed an Associate's Degree, 29% completed a Bachelor's degree, 7% completed a Master's degree, 18% completed a professional degree, and 6% chose not to respond.

Stimuli. Stimuli consisted of digitized high-definition video recordings of actors performing a series of actions. All infants saw a group-induction trial, three familiarization trials, and two test trials. A separate video was played for each trial. Each trial consisted of an initial phase followed by a final phase. The duration of the initial phase was fixed and identical for all participants. The duration of the final phase was infant-controlled. All trials are described from the infants' perspective.

Group-Induction Trial. All infants saw the same group-induction trial. At the start of the trial, three female actors sat around a table. Two of the actors (Fep-A, Fep-B) wore green t-shirts and green star-shaped glasses, while the third (Rab-A) wore a white button-up shirt and an orange sequined hat (see Figure 1). Actors that looked dissimilar on multiple features (i.e., different hair colors, hair styles, etc.) were selected to make it easier for infants to discriminate between them.

All actors began the trial with their heads down. During the 8-s initial phase of the trial, the actors looked at one another and labeled themselves: Fep-A said "Hi, I'm a Fep." Fep-B said, "Hi, I'm a Fep too." Finally, Rab-A said, "Hi, I'm a Rab." As each actor labeled herself, she looked back and forth between the other two actors. When not labeling themselves, the actors looked at the speaking actor as she spoke. After all actors had labeled themselves, the actors looked down and paused. The infants viewed this paused scene until the trial ended (see Apparatus and Procedure section for trial-ending criteria).

Familiarization trials. Infants then received three familiarization trials in which Fep-A demonstrated her social disposition (i.e., prosocial or antisocial) toward a neutral agent. The neutral agent looked very dissimilar to Fep-A and wore a plain red t-shirt. We used a neutral agent so that infants only had to keep track of one individual's social disposition. The valence of this disposition varied across infants.

In the *prosocial* valence condition, infants saw Fep-A behave prosocially three times. In the first familiarization trial, Fep-A and the neutral agent sat facing one another at a table. Fep-B and Rab-A were no longer present. Centered between them was a pile of multi-colored Legos. During the 15-s initial phase of the trial, the neutral agent selected and stacked four Legos of the same color. Fep-A built her own multi-colored stack while observing the neutral agent's actions. After stacking four blocks, the neutral agent looked for another Lego of the same color but was unable to find one. Then, Fep-A pointed to

the sought-after Lego so that neutral agent could locate it. Fep-A then placed the Lego on top her stack and smiled. The neutral agent then smiled at Fep-A and both agents proceeded to look down and pause until infants looked away (see Figure 2). The next two familiarization trials were highly similar except that on each trial the neutral agent built a different color stack of Legos, the neutral agent displayed a different seeking behavior, and hence Fep-A's prosocial responses varied. In the second familiarization trial, the neutral agent strained to reach a Lego but was unable to reach it; Fep-A pushed the block towards the neutral agent so that she could reach it. In the third familiarization trial, the neutral agent needed a Lego that was inside of a container but was unable to open it; Fep-A helped the neutral agent remove the lid from the container. Infants thus saw Fep-A behave prosocially three times.

The antisocial familiarization trials were identical to the prosocial familiarization trials except that Fep-A behaved antisocially by taking the Lego before the neutral agent located it (first trial), moving the Lego out of the agent's reach (second trial), and slamming the container closed when the agent tried to retrieve a Lego from it (third trial; see Figure 2). On each trial, upon Fep-A behaving antisocially, the neutral agent displayed sadness by furrowing her brow and pouting her lips while looking at Fep-A.

Same-agent test trials. Infants saw a prosocial and an antisocial test trial; the order of the test trials was counterbalanced across infants. In the *prosocial event*, Fep-A and the neutral agent again sat at a table seated across from one another with a pile of Legos between them (see Figure 3). During the 16-s initial phase of the trial, the neutral agent selected and stacked four Legos of the same color. Fep-A built her own multi-colored stack while observing the neutral agent's actions. After stacking four blocks, the neutral agent looked for another Lego of the same color. Once the agent located the one she wanted, she reached for the block but accidentally knocked it out of reach off of the table. Fep-A then retrieved the block and handed it to the neutral agent. The neutral agent then placed the block on her own tower, and smiled at Fep-A. Fep-A and the agent then looked down at the table and paused until the trial ended. The *antisocial event* was identical, except that instead of handing the block to the neutral agent, Fep-A put the block on her own tower and the neutral agent displayed sadness.

Different-agent test trials. The procedure for the different-agent condition was identical to that of the same-agent condition with one exception: in the test trials, infants saw a different Fep (Fep-B) than in the familiarization trials.

Apparatus and Procedure

Infants sat on their parent's lap 91.5 cm in front of a large television screen (68.5 cm x 122 cm). The room was dimly lit. A camera hidden at the base of the television (centered, 89 cm above the floor) recorded the infant's face during the experiment. Parents were instructed to close their eyes or look down to avoid biasing their infant's responses.

The television was connected to a Macintosh computer located to the left of the infant behind a sound-dampening room divider. This computer controlled the presentation of the experimental stimuli using custom software written in Python (Peirce, 2007). The software selected the correct version of each trial based on the infant's condition and presented the video in the center of the television screen (each video measured 64 cm x 37 cm on screen). The software also controlled the duration of each

trial. An experimenter observed the infant on a monitor and pressed a button on the keyboard whenever the infant attended to the video. The software separately computed looking times for the fixed-duration and infant-controlled portions of each trial; looking times during the infant-controlled portion of the trial were used to determine when each trial ended. In between trials, an attention-getter (a yellow smiley face measuring 28 cm x 20 cm) was displayed on the screen for 4 seconds and a brief tone was played to attract the infant's attention back to the television screen.

At the start of the experiment, the attention-getter was presented in the center of the television screen. When the infant attended to the screen, the experimenter initiated the presentation of the stimuli on the television screen. Infants first viewed the group induction trial, followed by the three familiarization trials appropriate for their condition. Each trial ended when the infant either (1) looked away for 2 consecutive seconds after having looked for at least 5 cumulative seconds or (2) looked for 60 cumulative seconds without looking away for at least 2 consecutive seconds.

Finally, infants viewed two test trials appropriate for their condition; half the infants in each condition saw the prosocial test trial first and half saw the antisocial test trial first. The test trials ended when infants (1) looked away for 2 consecutive seconds after having looked for at least 5 cumulative seconds or (2) looked for 30 cumulative seconds without looking away for at least 2 consecutive seconds.

Coding

In order to present events with trial duration contingent on the infant's attention, online coding was conducted by the experimenter (blind to condition and test trial), as described above. All infants were then coded offline from silent video by a trained coder who was naïve to the condition and the order of the test trials that the infant received. For each trial, the coder indicated the infant's direction of gaze (at the stimuli or away) for each frame of the video. Another trained, naïve coder coded all sessions offline. The two offline coders agreed on the child's direction of gaze for 96% of video frames. Trials in which agreement between the two coders was less than 90% (8/186) or the two coders disagreed on whether the child met the criteria to end the trial (1/186) were resolved by a third coder. With the exception of 6 trials in which the third coder agreed with the second coder (4 familiarization trials, 2 test trials), the primary coder's data were used in all analyses. In the cases where the third coder agreed with the second coder, the second coder's data was used instead.

The infants were highly attentive during the initial phase of the familiarization and test trials, attending, on average, for 98% and 90% of the initial phases, respectively. The main question of investigation was whether infants expected individuals to be consistent in their behavioral dispositions over time. Therefore, for the analyses, each test trial was categorized as either a consistent or inconsistent test event. In the consistent test event, Fep-A's actions in the test trial were consistent with her valence in the familiarization trials. In the inconsistent test event, Fep-A's actions differed in valence from her actions in the familiarization trials.

Preliminary analyses of the test data indicated no significant main effects or interactions with sex, all $Fs < 2.73$, $ps > .12$. The data were therefore collapsed across sex in subsequent analyses.

Results and discussion

Infants' looking times during the test trials were analyzed using an analysis of variance (ANOVA) with condition (same-agent, different-agent), familiarization valence (prosocial, antisocial), and test trial order as between-subjects factors and test event (inconsistent, consistent) as a within-subjects factor (See Figure 4). The analysis revealed a significant interaction of condition and test event, $F(1, 22) = 4.74$ $p = .04$, $\eta_p^2 = .18$. Planned comparisons indicated that infants in the same-agent condition looked reliably longer at the inconsistent ($M = 14.25$, $SE = 1.84$) than the consistent event ($M = 8.98$, $SE = 1.74$), $F(1, 22) = 5.06$, $p = .035$, $\eta_p^2 = .19$. In contrast, infants in the different-agent condition looked about equally to the inconsistent ($M = 13.30$, $SE = 1.84$) and consistent ($M = 15.25$, $SE = 1.74$) test events, $F < 1$

However, this two-way interaction was qualified by two three-way interactions. There was a marginal interaction of condition, trial order, and test event, $F(1, 22) = 3.64$ $p = .069$, $\eta_p^2 = .14$. There was also an interaction of condition, familiarization valence, and test event, $F(1, 22) = 4.74$ $p = .038$, $\eta_p^2 = .18$. No other effects were significant, all $Fs < 1.83$, all ps $> .19$.

Examination of the interaction of condition, trial order, and test event revealed that in the different-agent condition, infants looked longer at the second test event they were shown ($M_{Test1} = 13.3$, $SE_{Test1} = 1.84$; $M_{Test2} = 15.25$, $SE_{Test2} = 1.74$). This was not the case in the same-agent condition, where infants instead looked longer at the inconsistent event, regardless of order.

Exploration of the interaction of condition, familiarization valence, and test event indicated that in the different-agent condition, infants who received the antisocial familiarization trials looked longer at the inconsistent test event ($M = 16.23$, $SE = 2.67$) than at the consistent test event ($M = 14.56$, $SE = 2.56$). In contrast, infants who received the prosocial familiarization trials looked longer at the consistent test event ($M = 15.93$, $SE = 2.37$) than at the inconsistent test event ($M = 10.32$, $SE = 2.5$). Overall, this means that infants in the different-agent condition looked longer at the prosocial test event. This pattern of results did not emerge in the same-agent condition. These infants looked longer at the inconsistent test event ($M = 14.25$, $SE = 1.84$) than the consistent event ($M = 8.98$, $SE = 1.74$) regardless of familiarization valence.

In the same-agent condition, infants looked longer at the inconsistent test event, suggesting that infants expected Fep-A to be consistent in her behavioral dispositions. However, infants did not generalize Fep-A's behavioral disposition to Fep-B. It is possible that when infants in the different-agent condition saw the first test trial, they had no expectation of how Fep-B would behave because infants did not generalize Fep-A's behavioral disposition from the familiarization trials to the test trials. Instead, after seeing Fep-B behave in the first test trial, infants may have expected Fep-B to display the same behavioral disposition in the second test trial and found it unexpected if she did not. Thus, infants' looking time behavior may have been influenced by the order in which the test events occurred. In the different-agent condition, infants tended to look longer at the prosocial test events, suggesting that infants' looking time behavior was also influenced by the valence of the events. It appears that infants' reasoning about the stability of behavioral dispositions might be different for prosocial and antisocial individuals.

The interaction of condition, familiarization valence, and test event raises a potential issue given prior research that suggests that children react differently to positive and negative valence. For instance, Premack and Premack (1997) found that children tend to look longer at events that are prosocial in valence than antisocial in valence. By recoding the test events as inconsistent and consistent in the current analyses, the test events were collapsed across prosocial and antisocial events. If infants had a tendency to look longer at a particular valence, it may have been masked in the analyses.

To investigate this possibility, the test data was analyzed again in terms of the valence of the test events (prosocial, antisocial), rather than whether the test event was consistent or inconsistent with the familiarization trials. Infants' looking times during the test trial were then analyzed using an analysis of variance (ANOVA) with condition (same-agent, different-agent), familiarization valence (prosocial, antisocial), and test trial order as between-subjects factors and test event (prosocial, antisocial) as a within-subjects factor.

There was no main effect of event, $F < 1$, suggesting that infants did not have a preference for either the prosocial or antisocial event. There was a significant interaction of condition and event, $F(1, 22) = 4.86$ $p = .038$, $\eta_p^2 = .18$. This effect stemmed from the fact that infants in the same-agent condition looked slightly longer at the antisocial test event ($M = 13.43$, $SE = 1.90$) than the prosocial test event ($M = 9.80$, $SE = 1.67$), whereas those in the different-agent condition looked longer at the prosocial test event ($M = 16.11$, $SE = 1.67$) than the antisocial test event ($M = 12.44$, $SE = 1.90$). However, neither of these comparisons reached significance, both Fs < 2.46, ps > .13.

This interaction was qualified by a 3-way interaction of condition, familiarization valence, and test event, $F(1, 22) = 4.74$ $p = .04$, $\eta_p^2 = .18$. Conceptually, this is the same effect as the two-way interaction of condition and event reported in the previous set of analyses. In the same-agent condition, infants looked longer at the test event that was inconsistent with the valence they had seen in the familiarization trials, whereas those in the different-agent condition did not exhibit this pattern.

There was also a marginal 4-way interaction of condition, familiarization valence, test trial order, and test event, $F(1, 22) = 3.64$, $p = .069$, $\eta_p^2 = .14$. There were no other significant main effects or interactions, all $Fs < 1.83$, all ps > .19. This interaction emerged in part because infants' looking times in the different-agent condition were influenced by both trial order and test event valence, as discussed above. Thus, in the different-agent condition there was an overall tendency for infants to look longer at whichever test event they saw second. The one exception was infants in the different-agent condition that saw Fep-A behave prosocially and received the prosocial test event first. These infants looked longer at the prosocial event than the antisocial event. Because the majority of infants in the different-agent condition merely looked longer at the second test event they received regardless of its valence, there is no evidence that infants generalized behaviors across members of a social group.

In the same-agent condition, infants showed an overall pattern to look longer at the event that was inconsistent with how Fep-A behaved in the familiarization trials. The one exception was infants who saw Fep-A behave antisocially and received the antisocial test event first. These infants looked longer at the antisocial test event. Typically, in violation-of-expectation paradigms, there is a tendency for infants to look longer at the

first event they are shown. Because the antisocial event occurs first in this condition, infants' expectation of consistency may be overpowered by the order in which the test events occurred. If this is true, then perhaps with more analytic power, we would be able to detect infants' expectation of consistency above and beyond the effect of test event order. An alternative possibility is that there may be a difference in infants' attention to and/or reasoning about prosocial and antisocial individuals, and this difference in attention lead to the marginal 4-way interaction of condition, familiarization valence, test trial order, and test event. Perhaps infants treat prosocial and antisocial individuals as inherently different from one another, and this results in infants' ability to generalize Fep-A's prosocial behaviors across situations, but failure to generalize Fep-A's antisocial behaviors in some circumstances. If infants treat antisocial individuals differently than prosocial individuals, infants' ability to reason about antisocial behavioral dispositions may be more arduous and nuanced. Further investigation is needed to clarify these findings.

Overall, these results suggest there is a pattern for infants in the same-agent condition to expect a single agent to be consistent in her behaviors. If Fep-A was prosocial in the familiarization trials, infants expected her to display the same valence behavior in the test trial. In the different-agent condition, infants did not expect consistency in behaviors. If Fep-A was prosocial in the familiarization trials, infants had no expectations about whether a different Fep, Fep-B, would be prosocial or antisocial.

One possible interpretation of these findings is that infants did not generalize social disposition across members of a social group because social dispositions are not something infants readily attach to social group membership at this age. Alternatively, infants' ability to attribute dispositions and keep track of them across situations may be fragile at this age. It could be that any change in the situation could cause infants to have no expectation about how the person would behave in the test trial. The fact that infants who saw the antisocial familiarization trials and consistent test event first did not generalize Fep-A's behavioral disposition across situations, as discussed above, could be seen as evidence for this possibility. It may not have been that infants did not generalize characteristics across members of a social group, but that the presence of a new person in the test trial may have interfered with infants' ability to reason about social dispositions. If this possibility is true, then changing some other aspect of the situational context might cause infants to fail to generalize social dispositions even within the same agent.

Experiment 2

In Experiment 1, infants expected a single agent to be consistent in her social disposition. However, infants did not generalize social dispositions across members of a social group. Infants' failure to generalize behavioral dispositions in Experiment 1 may have occurred because infants do not readily attach behavioral dispositions to group membership. Alternatively, it is possible that result may have happened because infants' reasoning about behavioral dispositions is particularly fragile at this age. To test these possibilities, Experiment 2 investigated whether infants expect a single agent to remain consistent in her social disposition across situations. Infants saw the same familiarization trials as in Experiment 1, followed by test trials in which the same actor interacted with a neutral agent in a different context than in the familiarization trials (i.e. rolling a ball). If

infants expect a single agent to be consistent in her behavioral disposition across situations, it would suggest that in Experiment 1, infants' ability to encode social disposition information was sufficiently robust to support some forms of generalization. In turn, this would suggest that infants in in the different-agent condition of Experiment 1 did not generalize Fep-A's behavioral disposition to Fep-B because infants do not generalize social dispositions across members of a social group. If, however, infants do not expect a single agent to have a consistent behavioral disposition across situations, it would instead suggest that infants' failure to generalize across group membership in Experiment 1 was because infants' encoding of the social disposition was not sufficiently robust to support generalizations, even within agent.

Method

Participants. 15 healthy infants participated (7 female, 8 male; M_{age} = 19 months, 25 days, range 18 months, 10 days – 21 months, 22 days). Another 8 infants were tested but excluded because they failed to complete the experiment (6), because the child was highly active in 4 or more trials (1), or because the child was deaf (1). Seven infants were randomly assigned to the prosocial familiarization valence condition (M_{age} = 19 months, 25 days), and eight infants were randomly assigned to the antisocial familiarization valence condition (M_{age} = 19 months, 25 days).

Apparatus, Stimuli and Procedure

The apparatus, stimuli, and procedure were identical to Experiment 1 with the exception of the events shown in the test trial.

Infants saw two test trials. Infants received both a *prosocial event* and an *antisocial event* (order counterbalanced). In the *prosocial event*, Fep-A and the neutral agent sat at a table seated across from one another (see Figure 5). In the neutral agent's hand was a rubber ball. During the 13-s initial phase of the trial, the neutral agent raised the ball towards Fep-A to indicate that she wanted to play. Fep-A then nodded, and the neutral agent proceeded to roll the ball to her. Once Fep-A had the ball in her hands, she tossed the ball back-and-forth between her hands three times. Then, the neutral agent lifted her hands out towards the neutral agent indicating she wanted Fep-A to roll the ball back to her. Fep-A rolled the ball back to the agent. The neutral agent then tossed the ball back-and-forth between her hands three times. Then, she looked up at Fep-A and smiled. Fep-A and the agent then looked down at the table and paused until the trial ended.

In the *antisocial event*, once Fep-A had the ball, she tossed it back and forth between her hands. Then, Fep-A looked toward the agent, then back towards the ball and then proceeded to aggressively hide the ball in her own lap under the table. The neutral agent then looked down towards her hands, back up at Fep-A, and then cast her eyes downward and frowned. Fep-A and the agent then looked down at the table and paused until the trial ended (see Figure 5).

Coding

The coding was conducted in the same manner as Experiment 1. Trials in which agreement between the two coders was less than 90% (1/90) or the two coders disagreed on whether the child met the criteria to end the trial (2/90) were resolved by a third coder. With the exception of 1 test trial in which the third coder agreed with the second coder, the primary coder's data were used in all analyses. In the cases where the third coder agreed with the second coder, the second coder's data was used instead.

Results and discussion

Infants' looking times during the test trial were analyzed using an analysis of variance with familiarization valence (prosocial, antisocial) and test trial order as between-subjects factors and test event (consistent, inconsistent)[1] as a within-subjects factor.

There was a main effect of test event $F(1, 11) = 7.46$, $p = .02$, $\eta_p^2 = .40$. Planned comparisons revealed that infants looked reliably longer at the consistent test event ($M = 16.55$, $SE = 2.58$) than at the inconsistent test event ($M = 10.31$, $SE = 1.43$). This effect was qualified by a marginal interaction of familiarization valence and test event, $F(1, 11) = 4.11$, $p = .068$, $\eta_p^2 = .27$. There were no other significant main effects or interactions, all $Fs < 2.34$, all $ps > .15$. These results differ from the pattern that we saw in the same-agent condition in Experiment 1, suggesting that as we suspected, infants' understanding of behavioral dispositions as an enduring attribute is fragile at this age. Although this tendency to look longer at the consistent event is unusual for a violation-of-expectation paradigm, it is not without precedent. There are a small number of findings where children exhibit this looking time pattern (e.g., Kuhlmeier, Wynn, & Bloom, 2003). It is thought that when this pattern of looking time occurs, it is because the situation is especially challenging for children, and thus children may be preferentially focusing on the event that is easier to understand (Fiser & Aslin, 2002; Pelucchi, Hay, & Saffran, 2009). If this explanation is correct, it would be consistent with the idea that infants' understanding of behavioral dispositions is fragile at this age, resulting in infants preferentially attending to the event that is easier to process (i.e. the consistent event).

Exploration of the interaction of familiarization valence and test event revealed that when infants received the prosocial familiarization trials, they looked significantly longer at the consistent test event ($M = 21.54$, $SE = 3.73$) than at the inconsistent event ($M = 10.65$, $SE = 2.06$), $F(1, 11) = 10.84$, $p = .007$. When infants received the antisocial familiarization trials, there was a trend for infants to look longer at the consistent test event ($M = 11.57$, $SE = 3.57$) that at the inconsistent event ($M = 9.96$, $SE = 1.97$), but this difference did not reach significance, $F < 1$. However, it is not without precedent that infants responded differentially to the prosocial and antisocial valenced events given that two prior studies with 17-month-olds that investigated infants' expectations of stability in an individual's behavioral disposition also found differences in children's expectations depending on the valence of the event (Duh et al., 2017; Koch et al., 2015). In these studies, after seeing an individual behave antisocially, children expected them to behave antisocially again, but the same expectation of consistency did not occur when the individual was initially prosocial. Overall, these findings suggest that infants' understanding of prosocial and antisocial behaviors as enduring dispositions is particularly fragile at this young age.

General Discussion

Recent evidence suggests that infants expect members of a social group to share stable characteristics such as food preferences (Smith & Scott, 2017). However, older

[1] An analysis of the test data in terms of the valence of the test events (prosocial,

children expect that members of the same social group will share a variety of characteristics (Boseovski & Lee, 2006; Diesendruck & HaLevi, 2006; Martin et al., 1995; Shutts et al., 2013). In Experiment 1, infants were introduced to two arbitrary social groups, and then saw one member of a social group (Fep-A) either display a prosocial or antisocial disposition. In the test trial, infants either saw the same agent in the familiarization trial (Fep-A), or a different agent from the same social group (Fep-B), behave prosocially once and antisocially once. Infants expected Fep-A to be consistent in her behavioral disposition, and looked longer in the test trial if she was not. However, infants did not generalize behavioral dispositions across members of a social group. After watching Fep-A behave prosocially, infants had no expectation about how another Fep would behave. The current research provides the first empirical evidence that, at least under some circumstances, infants expect a single individual to be consistent in their behavioral dispositions.

Experiment 2 further clarified whether infants' failure to generalize behavioral dispositions in Experiment 1 was because: (1) infants do not attach social dispositions to social group membership at this age, or (2) infants have difficulty generalizing social dispositions. Specifically, Experiment 2 investigated whether infants expect a single agent to remain consistent in her social disposition across situations. After watching Fep-A behave prosocially in a lego-building context, infants did not expect the agent to be prosocial in a ball-rolling context, suggesting that infants did not expect a single agent to be consistent across situations. Because this pattern of results differs from the pattern seen in Experiment 1, it suggests that infants' failure to generalize in Experiment 1 was likely because infants' understanding of behavioral dispositions as enduring, stable attributes is fragile at this age. An additional indication of the fragility of infants' reasoning is that in some instances, infants' expectations of consistency were affected by valence. Specifically, in the different-agent condition of Experiment 1, and in Experiment 2, infants that received the antisocial familiarization trials displayed a different pattern of looking than did the infants that received the prosocial familiarization trials.

Experiment 1 adds important findings to the body of literature that investigates children's preferences for prosocial over antisocial others. In these studies, infants are placed in a forced-choice situation where infants are asked to pick between two agents (e.g., to look at object A or object B; to play with object A or object B) that either helped, or prevented someone from reaching their goal. In these situations, infants will typically choose to interact with the agent that was previously helpful (Hamlin, 2013; Hamlin & Wynn, 2011; Hamlin, Wynn, & Bloom, 2007; Hamlin, Wynn, & Bloom, 2010). This demonstrates that infants can clearly detect and distinguish between prosocial and antisocial individuals, and can use this information to make decisions about whom to interact with themselves. However, this does not necessarily mean that infants in these studies expect the agent's behavioral disposition to be enduring and consistent. Infants are only reasoning about the agent's behavior in a single instance, making it unclear whether infants truly expect agents to be behaviorally consistent. Experiment 1 provides evidence that within a given context, infants are able to abstract behavioral dispositions in a situation, and expect an agent to remain consistent in her behavioral dispositions.

However, at this age infants may be at a transitional stage. Infants may be just beginning to expect people to be consistent in their behavioral dispositions, and thus

require clear, concrete information in order to form such expectations. Infants' expectations of consistency can be disrupted by various factors such as the particular valence of an agent's behavior, or by a small change in the situational context. Changes in these factors may make the situation especially challenging for infants to reason about.

It is therefore possible that children's understanding of consistency in behavioral dispositions might benefit from additional supportive information. Perhaps infants need more instances of Fep-A's behavior in order to strengthen infants' expectations of consistency. It is also possible that additional, varied types of information might strengthen infants' expectations of consistency. In Experiment 1 and 2, Fep-A interacted with the same neutral agent in each familiarization trial. Therefore, infants may have interpreted Fep-A's behavioral disposition as pertaining to that particular individual. Perhaps if Fep-A interacted with a different neutral agent in each familiarization trial, it would demonstrate that Fep-A behaves a particular way regardless of the person they are interacting with, and infants would be more likely to form more abstract inferences about the stability of Fep-A's behavioral disposition (e.g., that Fep-A is a nice person). Providing stronger evidence may make it easier for infants to recognize stable dispositions. This could then lead infants to generalize behavioral dispositions within a single agent, across contexts.

The current studies provide no evidence that infants generalize dispositions across members of a social group. However, it could be that if provided with clearer information about the stability of the agent's behavioral disposition (as discussed above), infants may generalize behavioral dispositions across members of a social group. Alternatively, it is possible that at this age, infants do not attach behavioral dispositions to social group members. Thus, slightly older infants might demonstrate a more robust understanding of stable nature of dispositions, and expect that members of a social group will share behavioral dispositions.

Although infants in the current study did not generalize behavioral dispositions across social group members, prior research using a very similar paradigm to the one used in the current studies found that infants expect members of a social group to share food preferences (Smith & Scott, 2017). One possible reason this finding differed from the current results may be that food preferences are more robustly associated with group membership at this age than are behavioral dispositions. Previous work has argued that because food and eating are fundamentally social and cultural in nature (e.g., Liberman, Woodward, Sullivan, & Kinzler, 2016; Shutts, Kinzler, & DeJesus, 2013), infants may have a specialized, early emerging system for reasoning about foods. For example, Liberman et al. (2016) found that whereas infants generalized food preferences across members of a social group, infants did not generalize object preferences across members of a group. One possibility for this finding is that infants' reasoning about food may be qualitatively different than infants' reasoning about non-food items. As previously mentioned, reasoning about behavioral dispositions appears to be somewhat challenging for infants. Because infants may be particularly sensitive to food preferences as a marker of group membership, and because foods may be easier for infants to reason about, infants may associate food preferences with social groups at a younger age than behavioral dispositions.

Finally, these findings suggest there may be an important role for learning in infants' reasoning about behavioral dispositions. Whereas school-aged children and adults readily use group membership to make predictions about the behavior of others (e.g., Birnbaum et al., 2010; Bosekovski & Lee, 2006; Diesendruck & Ha Levi, 2006), infants in the current study did not. This may suggest that attributing characteristics to members of a social group may not be an automatic process. Instead, infants may first need to learn about the types of properties that they should associate with social group membership from their environment. One potential way children may learn about social groups may be through their parents. Previous research with 4-year-old children has found that the types of statements a parent makes when discussing a social group with their child can impact the beliefs and expectations that their child forms about that particular social category (Rhodes et al., 2012). Thus, it is possible that environmental influences such as parental input may influence children's social group based reasoning at even younger ages.

Chapter 3

Adults routinely use social group membership to draw inferences about the characteristics of others. Adults form essentialist beliefs about certain social categories, assuming that members of those social categories share inherent, stable characteristics. This makes social categories inductively powerful and can lead adults to use group membership to draw inferences about novel group members. Previous research suggests that the tendency to use social group membership to make predictions about the characteristics of others is robust by preschool. By this age, children form essentialist beliefs, assuming that social-category members share underlying essences that make them inherently similar to one another, but different form others (e.g., Bigler et al., 1997; Birnbaum et al., 2010; Diesendruck & HaLevi, 2006; Gelman et al., 1986; Martin et al., 1995; Rhodes & Mandalaywala, 2017; Shutts et al., 2013; Taylor, 1996; Waxman, 2013). Recent evidence suggests that the tendency to expect that members of a social group will be similar to one another emerges early in life. For example, infants as young as 20 months expect that members of the same social group, but not different social groups, will share stable characteristics such as food preferences (Smith & Scott, 2017).

Although this tendency may emerge early, the attributes children attend to when reasoning about social groups and the specific inferences that they make about them vary by age and culture. For example, when asked to make inductive inferences about the psychological or behavioral property of a novel character, children in Israel base their inferences on ethnicity information over other social categories (Birnbaum et al., 2010; Diesendruck & HaLevi, 2006). In contrast, children in Northern Ireland are more likely to use religion information as a basis of their inductive inferences (Smyth et al., 2017). This suggests that the particular inferences that children make about social groups, and the characteristics that children use when making these inferences must be learned.

It has been proposed that the manner in which social groups are discussed in the environment may impact how children form beliefs and expectations about social categories (Bigler & Liben, 2007). For example, highlighting an individual's group membership through explicit labeling or functional use of group labels (e.g., having girls line up for lunch first in the classroom) may influence children to believe that members of these groups share stable, non-obvious characteristics. Thus, even when adults distinguish among individuals in a completely neutral manner, children may infer that the grouping criteria is an important basis for categorization and construct beliefs about these groups based on labels. Additionally, after 5- and 9-year-olds either heard the phrases "Rosie eats carrots whenever she can" or "Rosie is a carrot-eater," the children who heard the latter judged that Rosie was more likely to eat carrots in the future (Gelman & Heyman, 1999).

Similar effects have been observed in children as young as four. Rhodes and colleagues found that 4-year-olds' beliefs about a social category were influenced by the types of statements they heard during a picture book session (Rhodes et al., 2012). An experimenter read a picture book about a novel social category, Zarpies, to the children. Each page presented a picture of a single person displaying a unique physical or behavioral property. A single line of text describing the depicted property accompanied each page using the language specified by the participant's condition. Some children

heard generic sentences about the social category (e.g., "Look at this Zarpie! Zarpies hate ice cream"), some heard specific sentences (e.g., "Look at this Zarpie! This one hates ice cream"), while some were provided with no social-group label (e.g., "Look at this one! This one hates ice cream"). Children who heard either type of generic statement (generic or specific) were more likely to endorse essentialist beliefs about Zarpies. That is, they were more likely to view Zarpie category members as sharing a deep, underlying inherent nature that make Zarpie members highly similar to one another, but distinctly different from other categories. Children who heard no social-group label were less likely to endorse essentialist beliefs about Zarpies.

These findings suggest that it is possible for an environmental factor as simple as the sentences used to describe a social category to affect children's expectations about social groups. However, in Rhodes et al. (2012), an experimenter read the book to the child multiple times over multiple research sessions. Therefore, it is unclear how children may learn about social groups through natural and more likely contexts, such as through how parents label and discuss social groups in brief, normal conversation. Further, the study used preschool-aged children who are known to form essentialist beliefs quite robustly. It remains unclear how children come to form these particular beliefs before essentialist tendencies are well established.

The current study explored one manner in which younger children may learn about social groups in their environment: parental input. Three-year-olds and their parents first viewed a picture book about a novel social category, Zarpies, adapted from Rhodes et al. (2012). Across participants, we varied the types of sentences included in book. Half of the participants received the generic book sentences, and half received the no-label book sentences. Then, in a violation-of-expectation video task, children saw a Zarpie display behaviors that were either consistent or inconsistent with those in the picture book. We then examined whether children learned about the Zarpies characteristics through the brief parental action and expected the Zarpies in the video task to behave similarly. Additionally, we examined whether specific types of parental input predicted children's learning about the Zarpie characteristics.

We predicted that children who received the generic book sentences would be more likely to learn about the Zarpie behaviors in the book and look longer at the Zarpie-inconsistent events than children who received the no-label book sentences. Further, we predicted that parents who talked more about the Zarpie behaviors and highlighted similarities and differences between the child and Zarpies would display more robust learning about the Zarpie behaviors.

Experiment 3

Method

Participants. 63 healthy children participated (33 female, 30 male; M_{age}= 36 months, 17 days, range 35 months, 6 days – 39 months, 9 days). Another 4 children were tested but excluded because they failed to complete the picture-book task (2), because they failed to complete the video task (1), or because they had an insufficient number of useable trials in the video task (1)[2]. Approximately equal numbers of children were

[2] See Coding section for information on eliminated trials.

randomly assigned to either the generic statement condition (N = 32) or the no-label statement condition (N = 31).

The children's names in this and the following experiment were obtained from birth records provided by the California Department of Public Health, as well as from a database of parents who had previously expressed interest in participating in research studies with their children. Parents were offered reimbursement for their transportation expenses, and their child was given a small gift (book or t-shirt) for participating. Parents provided written informed consent for their child's participation.

The racial and ethnic composition of the final sample was 64% Caucasian, 2% Asian, 3% American Indian or Alaska Native, 6% Black or African-American, 2% Native Hawaiian or Pacific Islander; an additional 6% chose 'other race', 11% selected more than one race, and 6% chose not to respond. 41% of the sample identified as Hispanic or Latino, 53% identified as not Hispanic or Latino, and 6% chose not respond. In lieu of income information, we recorded the highest level of education reported by either parent: 2% completed less than high school, 17% completed high school, 28% completed an Associate's Degree, 30% completed a Bachelor's degree, 5% completed a Master's degree, 16% completed a professional degree, and 2% chose not to respond.

Apparatus

The picture-book task was conducted in one room and the video task took place in a different room.

Picture-book task. Children sat on their parent's lap facing a table. A camera was centered on the table to capture the parents' verbal utterances. A second camera was mounted above and behind the child in order to capture the stimuli.

Video-task. Children sat on their parent's lap 91.5 cm in front of a large television screen (68.5 cm x 122 cm). The room was dimly lit. A camera hidden at the base of the television (centered, 89 cm above the floor) recorded the child's face during the experiment. The television was connected to a Macintosh computer located to the left of the child behind a sound-dampening room divider. This computer controlled the presentation of the experimental stimuli using custom software written in Python (Peirce, 2007). The software selected the correct version of each trial based on the child's condition and presented the video in the center of the television screen (each video measured 64 cm x 37 cm on screen).

Stimuli

Picture-book Task. The stimuli used in the picture-book task were adapted from Rhodes and colleagues (2012). The book consisted of 16 illustrated pages (identical across conditions; see Figure 6 for example pages). Each page featured a person displaying a unique characteristic or property (see Table 1 for the full list of properties). The characters were diverse in terms of sex, ethnicity and age. Although no single clothing item was a defining feature of the characters, the characters wore category-typical clothing that allowed for visual identification of the category. Each illustration was preceded by a page that had a statement that described the illustration. This statement varied based on the participant's condition. In the generic statement condition, each statement about the novel social group was a generic statement (e.g., "Look at this Zarpie! Zarpies hate ice cream"). In the no-label condition, the statement described the

action on the page, but did not label the character as belonging to a social category (e.g., "Look at this one! This one hates ice cream").

The stimuli in Experiment 3 also included a paragraph designed to introduce parents to the content of the book. This paragraph was adapted from Experiment 3 of Rhodes and colleagues (2012) and was intended to induce parents to hold essentialist beliefs about Zarpies (see Appendix A). In this paragraph, Zarpie group members were described as highly similar to one another in terms of biology and cultural practices, but very different from other groups' characteristics.

Video Task. Stimuli consisted of digitized high-definition video recordings of actors performing a series of actions. In each trial, children saw an agent wearing an outfit similar to the characters featured in the picture-book task (i.e., khaki pants, a yellow long-sleeved shirt, and a blue hat). The videos featured four different agents; half of the agents were male, half were female. All children saw six test trials. A separate video was played for each trial. Each trial consisted of an initial phase followed by a final phase. The duration of the initial phase was fixed and identical for all participants. The duration of the final phase was child-controlled.

Trials were arranged in three pairs. Each pair focused on one of the behaviors shown in the book. In one trial in each pair, the agent showed a behavior that was consistent with the Zarpie behavior depicted in the book. In the other trial in each pair, the agent displayed a behavior that was inconsistent with the Zarpie behavior depicted in the book (see Figure 7). All trials are described from the children's perspective.

In the first pair of test trials, children saw the Zarpie character either bounce a ball on their head (consistent-ball event) or bounce a ball on their knee (inconsistent event). During the 9-s initial phase of the consistent event, the male agent stood facing the camera, holding the ball out in front of him. He then raised the ball to his head and bounced the ball on his head twice, catching the ball each time. Then, with the ball in both hands he returned the ball to its initial starting position and paused. The children viewed this paused scene until the trial ended (see Procedure section for trial-ending criteria). The inconsistent event was identical, except that the Zarpie character bounced the ball on his knee instead of his head.

In the second pair of trials, children saw the Zarpie character either hate ice cream (consistent event) or like ice cream (inconsistent event). During the 8-s initial phase of the consistent event, a female agent stood facing the camera. Positioned to her left was an ice cream cone on a small table. She reached over and grabbed the ice cream, raised it to her face, and pretended to taste the ice cream. Then, she moved the ice cream away from her while saying "Yuck!" and displaying negative affect to indicate that she disliked the ice cream. After repeating these actions once more, she placed the ice cream back on the table in its initial position, returned to her starting position, and paused. The children viewed this paused scene until the trial ended. The inconsistent event was identical, except that the Zarpie character said "Mmm!" and smiled displaying positive affect while tasting the ice cream.

In the third pair of test trials, children saw the Zarpie character either flap their arms while happy (consistent event) or spin in circles (inconsistent event). During the 6-s initial phase of the consistent event, a male agent stood facing the camera. He then smiled, and with his arms placed outward, and moved his arms up and down twice. He

then paused with his arms at his side for 1-s and proceeded to repeat the arm flapping action. He then returned to his initial starting position and paused until the trial ended. The inconsistent event was identical except the male agent spun in circles instead of flapping his arms. The male agent spun in a circle twice, paused in a neutral position for 1-s, and then repeated the spinning action. He then returned to a neutral position and paused.

Procedure

Children played with toys in the lobby while their parent completed the consent form and MCDI-III. After completing the paperwork, parents were given the essentialist-belief inducing paragraph to read. Immediately after the parent finished reading the paragraph, the parent and child were brought into an adjoining room where the child sat on their parent's lap and completed the picture-book task. Immediately following the picture-book task, children went into a different adjoining room and participated in the video task. The duration of the entire study was approximately 30 minutes.

Picture-book task. Parents were asked to read the statements aloud and describe the pictures to their child as they would at home. The experimenter then left the parent and child in the room alone to complete the task. Parents were instructed to return to the lobby when they were finished.

Video task. Immediately following the picture book task, children participated in a video task. At the start of the experiment, an attention-getter, a yellow smiley face measuring 28 x 20 cm, was presented in the center of the television screen. When the child attended to the screen, the experimenter initiated the presentation of the stimuli on the television screen. Children viewed the three pairs of test trials appropriate for their condition; half the children in each condition saw the consistent trial first and half saw an inconsistent test trial first. Each test trial ended when the child either (1) looked away for 1 consecutive second after having looked for at least 5 cumulative seconds or (2) looked for 60 cumulative seconds without looking away for at least 1 consecutive second.

After each trial, an attention-getting stimulus (a yellow smiley face measuring 28 cm x 20 cm) was displayed on the screen for 4 seconds and a brief tone was played to attract the child's attention back to the television screen. After trial 4, the smiley-face was replaced by an image of a baby (measuring 30 x 19cm) that was accompanied by a 4-s audio clip of a baby laughing.

Coding

Picture-book Task. The picture-book task was coded and transcribed by naïve research assistants. Parent's language was first transcribed verbatim. Non-verbal sounds or exclamations were not included in the transcription (e.g. gasps, sighs, groans, etc.). Children's language was not transcribed.

Following transcription, parent utterances were coded in several ways. First, we coded whether the utterance referred to a Zarpie characteristic and, if so, which characteristic it referred to. For each utterance about a Zarpie characteristic, we next coded whether the utterance fell into any of the following categories: generic, comparative, and contrastive. Generic statements were any statements that broadly described the category as a whole or made generalized statements about the behavior or characteristic, rather than describing particular group members. For example, "Those Zarpies hate ice cream" would be coded as a generic statement. Statements that

highlighted how the child displayed similar characteristics or behaviors to the characters in the book (e.g., "You like to flap your arms too, don't you?") were coded as comparative statements. Statements that highlighted how the child was different than the characters in the book (e.g., "You don't hate ice cream, you love ice cream!") were coded as contrastive statements.

A second individual transcribed and coded 25% of the picture-book sessions. Cohen's κ was calculated to determine agreement between the two coders. There was high agreement between the two coders for all categories of statements, all $\kappa > .81$, all $p < .001$.

In order to control for differences in parent verbosity, parental statements about each characteristic displayed in the book (e.g., bouncing a ball) were converted to a percentage by taking the raw number of parental statements about that characteristic and dividing by the total number of parental utterances. To capture parents' general tendency to engage in a particular type of talk, generic, contrastive, and comparative statements were also converted to a percentage by taking the raw number of the parental statements in a particular statement-type category (e.g., generic statements), and dividing by the total number of parental statements.

Video Task. In order to present events with trial duration contingent on the children's attention, online coding was conducted by the experimenter (blind to condition and test trial), as described above. All children were then coded offline from silent video by a trained coder who was naïve to the condition and the order of the test trials that the child received; the looking times resulting from this coding were used in all analyses. For each trial, the coder indicated the child's direction of gaze (at the stimuli or away) for each frame of the video. Another trained coder who was naïve to the child's condition and order of the test trials coded all sessions. The two offline coder's agreed on the child's direction of gaze for 96% of video frames. Trials in which agreement between the two coders was less than 90% (22/186) or the two coders disagreed on whether the child met the criteria to end the trial (7/186) were resolved by a third coder. With the exception of 12 trials in which the third coder agreed with the second coder, the primary coder's data were used in all analyses.

For some trials, both offline coders agreed that the online coder terminated the trial prematurely before the child met the looking criteria (25 out of 378 trials). When this occurred, both trials for the affected item (i.e. ball, ice cream, or flap/spin) were eliminated from the analyses. Individual items were also eliminated if the difference between participants' looking times to the inconsistent and consistent events was more than 2.5 standard deviations away from the mean of the condition (4 items). Finally, one item was excluded because the participant looked the maximum amount of time in both trials. Children were excluded if they failed to contribute at least two useable items to the analysis. The final sample consisted of 26 ball items, 30 ice cream items, and 25 flap/spin items in the generic condition and 26 ball items, 30 ice cream items and 25 flap/spin items in the no-label condition.

Children were highly attentive during the initial phase of each test trial, attending, on average, for 95% (ball trials), 98% (ice cream trials), and 93% (flap/spin trials) of the initial phases, respectively. Children's looking times to final paused scenes were calculated for each event. We then calculated a difference-score for each item by

subtracting children's looking time to the consistent event from their looking time to the inconsistent event for that item. The difference scores were then averaged across all three items to create an overall difference score for each participant.

Preliminary analyses of participants' overall difference scores indicated no significant interactions of sex or order with condition, all $Fs < 1.82$, $ps > .18$. The data were therefore collapsed across these factors in subsequent analyses.

Results and Discussion

Children's overall difference scores were analyzed using an analysis of variance (ANOVA) with condition (generic, no-label) as a between-subjects factor. There was a main effect of condition, $F(1, 61) = 6.93$, $p = .011$, $\eta_p^2 = .10$. One-sample t-tests revealed that in the generic condition, children's overall difference scores were significantly different from chance (0) ($M = 3.49$, $SD = 6.42$), $t(31) = 3.08$, $p = .004$, $d = 1.11$. In the non-generic condition, children's overall difference scores were not significantly different from chance ($M = -0.4$, $SD = 5.24$), $t < 1$. As shown in Figure 8, in the generic condition, all of the items had positive difference scores, indicating that children looked longer at the Zarpie-inconsistent events. In contrast, no such pattern emerged in the non-generic condition.

These results suggest that children in the generic condition learned something about how Zarpies typically behave from a brief interaction with their parent, and expected that a novel Zarpie in the video task would behave in a manner that was consistent with the Zarpie behavior depicted in the book. This pattern of learning was not present in the no-label condition. Overall, these results suggest that exposure to generic sentences increased children's tendency to view Zarpies a coherent group that had members that were highly similar to one another.

In the next set of analyses we investigated whether, above and beyond the sentences in the book, any other dimensions of parents' talk differed across conditions. The percentage of parental utterances about each item in the video task (ball, ice cream, flap/spin) were entered into a multivariate analysis of variance (MANOVA) with condition (generic, no-label) as a between-subjects factor (see Table 2). There was no effect of condition, $F < 1$, suggesting that the amount of talk about the specific items did not differ across conditions.

We next examined whether the way in which parents talked about Zarpies differed across conditions. Independent-samples t-tests revealed that the percentage of utterances containing generic statements (excluding the generic statements that were written in the book) was significantly higher in the generic condition ($M = 1.49$, $SD = 1.89$) than in the no-label condition ($M = .24$, $SD = 1.02$), $t(61) = 3.27$, $p = .002$, $d = .82$. Similarly, the percentage of parental utterances containing comparative statements was significantly higher in the generic condition ($M = 1.90$, $SD = 1.90$) than in the no-label condition ($M = .85$, $SD = 1.26$), $t(61) = 2.58$, $p = .012$, $d = .65$. The percentage of parental utterances containing contrastive statements was also marginally higher in the generic condition ($M = 2.36$, $SD = 3.23$) than in the no-label condition ($M = 1.15$, $SD = 1.81$), $t(61) = 1.83$, $p = .073$, $d = .46$. Together, these findings suggest that parents who received the generic statements in the picture-book task spontaneously produced more generic statements, above and beyond the book sentences that were provided to them. These

parents also highlighted similarities and differences between the child and the Zarpies more often than parents who did not receive the generic statements.

Next, we examined whether individual differences in parental talk contributed to children's performance in the video task using a series of multi-level mixed-effect models. The lmer function of the lme4 package (Bates, Maechler, Bolker & Walker, 2015) was utilized using the statistical software RStudio (R Core Team, 2015). The children's difference scores for each test item were entered into two multi-level mixed-effects models with random effects for subjects (Baayen, Davidson & Bates, 2008). The first model represented the Null model and did not contain the fixed effect of interest. The second model was identical except that it included a fixed factor of interest. The two models were compared using -2 log-likelihood ratio tests.

We analyzed children's difference scores with 4 different models, each of which included one of the following parent-talk scores as a fixed effect: (1) percentage of utterances that contained talk about each item in the video task, (2) the percentage utterances that contained generic statements, (3) the percentage of utterances that contained comparative statements, and (4) the percentage of utterances that contained contrastive statements. In each case, the model was compared to a null model without the parental talk variable of interest. No significant effects were found, all $\chi^2 < 1$, $ps > .5$. These findings suggest that these types of parental talk did not influence children's performance in the video task, above and beyond the sentences included in the book. However, many parents made these types of statements infrequently, or not at all, and thus there were many zero scores for these categories. It is possible that with a larger sample, or a larger sample of talk from each parent, that we would see significant effects of these types of parent talk.

General Discussion

The current study examined how children may learn about social groups in their environment through one proximal environmental influence, parental talk. Parents first read an essentialist-belief inducing paragraph that introduced them to the novel social group called Zarpies, and discussed how Zarpies members had very similar characteristics, but very different characteristics from members of other social groups. Parents then read a picturebook to their child that depicted Zarpie characters engaging in a variety of activities; the book either contained generic sentences or no generic sentences. Following the picture book task, children participated in a video task where they were shown individuals dressed as Zarpies display behaviors that were either consistent or inconsistent with the Zarpie behaviors depicted in the book. Children that received the generic sentences in the book expected the Zarpie characters in the video task to behave in a Zarpie-consistent manner, and looked longer in the video task if the Zarpie characters displayed Zarpie-inconsistent behaviors instead. This pattern of learning was not observed when children received the no-label sentences: children instead looked equally to the Zarpie-consistent and -inconsistent behaviors.

These results provide the first evidence that toddlers can learn about social groups present in their environment from a brief parental interaction. After parents read the 16 generic statements provided in the book to their child, children in the generic statement condition formed an expectation that Zarpie members would be similar to one another

and display the same types of behaviors. This suggests that, from little environmental input, children can learn which groups are likely to be essentialized. That is, they can learn which groups will likely have individuals that share stable, inherent characteristics that make individuals in the group similar to one another, but different than other groups. Children can then use this information to guide their expectations about the characteristics of social groups.

These findings also provide the first evidence that generic statements can impact the beliefs and expectations that children form about social categories at even younger ages than have previously been examined. Although previous studies have found that parents produce generic statements with children as young as 2 in reference to animal and artifact categories (Gelman, Coley, Rosengren, Hartman, & Pappas, 1998; Gelman, Ware, Kleinberg, Manczak, & Stilwell, 2014; Pappas & Gelman, 1998), research has yet to examine how exposing children to these statements directly impacts children's beliefs and expectations about social categories.

The fact that hearing these generic statements impacted the beliefs children formed about Zarpies is especially impressive given that generic statements are a sophisticated, abstract aspect of language. Because generic statements can be expressed in many ways, the linguistic markers of generic statements are rather ambiguous. For instance, "The dog likes to dig", "Dogs play fetch", and "A dog is a man's best friend" are all examples of generic statements. Yet, highly similar statements such as, "The dog dug up the backyard", "Dogs are playing fetch in the park", and "A dog ate my homework" are non-generic statements. Thus, identifying a generic statement, and its potential meaning, requires some understanding of morphosyntactic cues (Gelman, 2004; Gelman & Raman, 2003). In addition, generic statements are conceptually abstract. For example, it is impossible to observe an entire kind (e.g., dogs) in the environment. Instead, one can only observe particular exemplars of a kind (e.g., a dog/dogs). Thus, in order to determine what a generic statement refers to, one must make an inferential leap (Gelman, Goetz, Sarnecka, & Flukes, 2008). Despite these complexities, results from the current study suggest that young children detect this type of language in their environment, and that these statements have a powerful impact on children's social-group based reasoning.

Children's understanding and production of generic statements from a young age may suggest that thinking in terms of categories may be privileged in children's thinking about the social world. Although the social world can be carved up in an infinite amount of ways, this linguistic cue may signal to children meaningful categories in the environment, and cause children to attend to the features of the category that may mark group membership. Children can then use this information to make inductive inferences about the likely properties of a group.

In the current study, all parents were given the same essentialist-belief inducing paragraph. Yet, it is clear that the presence of the generic statements in the generic statement condition affected the parents' language and behaviors in the picture-book task. Parents in the generic statement condition produced more generic (above and beyond the statements provided), comparative, and contrastive statements than parents in the no-label condition. Parents in both picture-book conditions talked about each of the items in the video task equally. Thus, it is not the case that parents in the generic condition merely

produced more statements about a particular item. Instead, it may be that the mere presence of the generic statements in the picture book caused parents to spontaneously hold stronger essentialist beliefs than parents in the no-label condition, and thus produce more statements that highlighted how Zarpie members a part of a unique, coherent group whose members are highly similar to one another. These findings support theoretical claims that subtle environmental factors, such as functional use of group labels, implicit or explicit labeling (Bigler & Liben, 2007), or even the statements in a children's book, may influence the particular beliefs that parents form about a particular social group, in turn impacting the type of language their child hears and the social-group based beliefs their child forms.

However, the current study did not find that other types of parental input (i.e., generic, comparative, and contrastive statements) predicted children's learning about the Zarpie characteristics. One reason for this may be that parents produced these types of statements in low percentages, or in some cases not at all. Previous research has found that in picture-book reading contexts, up to 3% of parental utterances contained generic statements (Gelman & Raman, 2003). Although this is a larger percentage than observed in the current study, it is still a relatively low percentage. It is possible that these other types of statements do in fact influence children's social-group based beliefs, but in order to detect these effects we would need more observations.

In summary, the findings from the current study suggest that children can learn about the characteristics of a novel social group from brief parental interaction, and form expectations about the characteristics of social group members based on this input. These findings highlight the powerful impact local environmental influences have on the formation of children's social-group based beliefs.

Chapter 4

By preschool age, children demonstrate essentialist beliefs about social categories: children believe members of a social category will share deep, inherent natures and that members of a category will be similar to one another in both obvious and nonobvious ways. This can lead to generalizing prior knowledge about a category to novel entities and situations (Rhodes & Mandalaywala, 2017). For example, by age 4, American children essentialize gender, predicting that a baby born as a boy will inevitably grow up to demonstrate male-typical characteristics (e.g., liking football), despite environmental influences (e.g., growing up on an island of women that liked ballet; Taylor, 1996; Taylor et al., 2009).

Although the tendency to use social-group membership to predict others' characteristics is well established by preschool (e.g., Gelman, 2003), the specific inferences that children make about social groups, and the attributes children attend to when making these inferences, vary with age and culture (for a review, see Rhodes & Mandalaywala, 2017). For example, essentialist beliefs about gender in US children living in a liberal community declined with age, whereas essentialist beliefs about gender did not decline in children from a more conservative community (Rhodes & Gelman, 2009). Additionally, 5- and 7-year-old White children from both urban and rural areas of the US do not initially use race as a basis to draw inferences about the characteristics of others. Instead, children gradually begin to use race as a basis for inference as they get older (Rhodes & Gelman, 2009).

Although previous studies have focused on individual social categories (e.g., Diesendruck & Halevi, 2006; Rhodes & Gelman, 2009), typically people belong to multiple social categories simultaneously (e.g., a Christian female). Recent evidence from studies using forced-choice inferences tasks suggest that the particular categories that children attend to when an individual belongs to multiple social groups vary with culture as well. In these tasks, an experimenter first introduces children to two test characters, each of which belongs to two social groups (e.g., an Arabic boy and a Jewish girl). The experimenter introduces each test character, labels both of the characters' social groups, and provides a novel fact about the character (e.g., "This is a boy. He is Arabic. He wants to be a Mashitz when he grows up; "This is a girl. She is Jewish. She wants to be a Nagim when she grows up."). Then, the experimenter introduces a target character that matches one social category of each test character (e.g., "Look at this girl. She is a girl like this girl. She is Arabic like this boy"). After labeling the target character's social categories, the experimenter asks the child which test character the target is most similar to.

In one recent study, Birnbaum and colleagues (2010) used this force-choice inference task to examine the inductive inferences made by 5- to 11-year-old children in Israel (Birnbaum, et al., 2010). The task involved three social categories: ethnicity (Jewish, Arabic), religiosity (religious, secular) and gender (boy, girl). Children tended to select the test character that matched the target character's ethnicity as opposed to the target character's religiosity or gender. This suggests that in Israel, children are more likely to base their inferences on ethnicity information than on religion or gender information. This was especially true of religious Jewish children, who chose the ethnicity match at higher rates than either secular Jewish children or Muslim Arab

children in Israel. In contrast, Smyth and colleagues (2017) found that when 6- to 11-year-olds in Northern Ireland were tested with a similar forced-choice inference task, they preferentially based their inferences on religion (Catholic/Protestant) over other social categories (Smyth et al., 2017).

The contrast in findings between these two studies demonstrates that culture influences the particular social categories a child uses when making inductive inferences about the characteristics of a novel individual. One reason that children in Israel may base their inductive inferences on ethnicity over other social categories is because categories tied to ethnicity are highly salient in Israel. Ethnicity categories such as "Arab" and "Jewish" may be especially salient to Israeli children because these categories are unavoidably tied to historic, violent conflicts that have happened in Israel. These ethnic categories are often highlighted in Israeli media and school materials (Bar-Tal & Teichman, 2005). Similarly, social categories based on religion are particularly salient to children in Northern Ireland due to historic, violent political conflicts dating back centuries. To this day, membership in a particular religious category such as "Catholic" or "Protestant" is a critically important dimension in Northern Ireland (Gillespie, 2010), and Northern Irish communities are subject to residential (Lloyd & Shuttleworth, 2012), martial (Lloyd & Robinson, 2011), and educational segregation (Gallagher, 2010) based on religious affiliation.

Further evidence suggests that even within a given culture, local differences in the contexts that children experience can impact the inferences that they draw about social categories. Children in Israel who attend bicultural schools (half of the students in each class are Arabic and half are Jewish) are less likely to make inductive inferences based on ethnicity than children who attend mono-cultural schools that are predominantly one ethnicity (Deeb, Segall, Birnbaum, Ben-Eliyahu, & Diesendruck, 2011; Smyth et al., 2017). Smyth and colleagues (2017) found a similar pattern in Northern Ireland: children attending diverse religious schools with both Protestant and Catholic children were less likely to base their inferences on religion than children attending mono-religion schools. Together, these findings suggest that exposure to social diversity, and intergroup contact where children regularly encounter members of contrasting social categories, such as multiple different religions, decreases the likelihood that children will use that particular social category as a basis of inductive inference.

The environment of children growing up in the United States (US) differs from that of children in Israel and Ireland in several ways that might influence children's social-group based inferences. First, in the United States, social categories based on ethnicity and religion are not as historically charged as in Israel and Northern Ireland, respectively, and thus it is likely that these categories are less salient to American children. Second, many children in the US experience environments that are diverse with respect to religion and ethnicity. These environmental contexts may be more reflective of the integrated schools in Israel and Ireland. This leads to the prediction that children in the US should be unlikely to use religion and ethnicity as a basis for inductive inferences about others' characteristics.

Yet, few studies have investigated the relative inductive potential of social categories in the US, and the studies that have investigated this question have focused largely on gender and race (e.g., Shutts et al., 2013). Thus, the extent to which children in

the US use ethnicity and religion to make inductive inferences remains unclear. To my knowledge, there has been no research on how children in the US reason about ethnicity, and only one study that investigated how US children reason about religion (Smyth et al., 2017). This study found that 6- to 10-year-old children in the Boston area were no more likely to base inductive inferences on religion (Protestant/Catholic) than on pet ownership (goldfish/hamster owner). This study provides preliminary support for the prediction that children in the US do not readily make inferences based on religion.

However, this small amount of research on children's reasoning about religion categories, as well as the studies investigating children's reasoning about race and gender categories in the US, have tended to focus on upper-middle class White urban samples (Shutts et al., 2013; Smyth et al., 2017). The findings described above show that even within a country, children's inferences can differ based on local environmental context (e.g., Deeb et al., 2011; Smyth et al., 2017). Similarly, even within the US, children's beliefs about gender have been shown to differ across rural and urban environments (Rhodes & Gelman, 2009). It is therefore important to examine how children use social categories to make inductive inferences in a different sample that is more ethnically, economically, and culturally diverse.

The goal of the present study was thus to examine how US children growing up in a diverse context reason about ethnicity and religion. In order to facilitate comparisons with other studies conducted in the US, we also investigated how children reason about gender. To address this question, 5- to 6-year-old children were tested with a forced-choice inference task adapted from Birnbaum et al. (2010). This age sample was chosen to be similar to the previous studies that have been conducted on this topic, and to investigate whether children's inferences change as a function of age.

The current sample was recruited from the San Joaquin Valley (a rural, agricultural region with high levels of cultural, ethnic, and economic diversity). Thus, our sample is more culturally, ethnically, and economically diverse than prior studies that have investigated this topic in the US. More importantly, this area tends not to be highly segregated. Prior findings that children in religiously or ethnically integrated schools pay less attention to religion (Smyth et al., 2017) and ethnicity (Birnbaum et al., 2010; Deeb et al., 2011), respectively, suggests that being in an integrated environment may reduce one's attention to a particular social dimension. Therefore, children in our sample may be less likely to attend to categories such as religion and ethnicity. Additionally, prior findings suggest that children in the US attend to gender from an early age (Rhodes & Gelman, 2009). In our study, we expect that children will be less likely to attend to religion and ethnicity than children in Northern Ireland or Israel, and instead will robustly attend to gender when making inductive inferences. However, because children's reasoning about gender and race changes with age (Rhodes & Gelman, 2009), it is possible that these patterns will change as children become older.

Experiment 4

Method

Participants. 32 5-year-old children (57.30 — 71.72 months; M = 66.34 months; 16 female) and 32 6-year-old children (72.1 — 83.9 months, M = 78.21 months; 16 female) participated in the study. This sample size was selected based on previous studies using a similar method (i.e. Birnbaum et al., 2010). An additional 10 children were tested

(8 5-year-olds, 2 6-year-olds), but were excluded due to an experimenter error (1), because the child asked to stop participating (1), or because they were missing a response on at least 1 trial[3] (8).

The children's names were obtained from a database of parents who had previously expressed interest in participating in research studies with their children. Parents were offered reimbursement for their transportation expenses, and their child was given a book for participating. Parents provided written informed consent for their child's participation. The Institutional Review Board at University of California, Merced approved all procedures.

The racial and ethnic composition of the 5-year-old participants in the final sample was 72% Caucasian, and 3% American Indian or Alaska Native. An additional 13% chose 'other race', 9% selected more than one race, and 3% chose not to respond. 50% of the sample identified as Hispanic or Latino, 47% identified as not Hispanic or Latino, and 3% chose not respond. In lieu of income information, we recorded the highest level of education reported by either parent: 19% completed high school, 16% completed an Associate's Degree, 31% completed a Bachelor's degree, 19% completed a Master's degree, and 16% completed a professional degree.

The racial and ethnic composition of the 6-year-old participants in the final sample was 47% Caucasian, 6% Asian, 3% American Indian or Alaska Native, 6% Black or African-American; an additional 25% chose 'other race', 7% selected more than one race, and 6% chose not to respond. 59% of the sample identified as Hispanic or Latino, 35% identified as not Hispanic or Latino, and 6% chose not respond. Thirty-seven percent of parents completed high school, 16% completed an Associate's Degree, 28% completed a Bachelor's degree, 13% completed a Master's degree, and 6% completed a professional degree.

Stimuli. Stimuli included 36 pictures taken from the "Child Affective Facial Expression (CAFE) set" by LoBue and Thrasher (2015). The pictures showed 4- to 6-year-old children's faces displaying happy expressions (see Figure 9). Each picture was 3" x 3", printed on cardstock, and laminated. The stimuli varied in gender and ethnicity. Half of the stimuli were male and half were female. Half of the stimuli were Hispanic and half were non-Hispanic.

Apparatus and Procedure. The duration of the entire study was approximately 30 minutes. Children played with toys in the lobby while their parent completed the consent form and a vocabulary checklist (see Appendix B). The vocabulary checklist included ethnicity (White, Latino, Mexican, Hispanic) and religion words (Protestant, Catholic, Christian, Jewish). Parents were asked to indicate whether their child knew any of the words.

The children and their parents were then brought into an adjoining room. Children were seated at a 42" x 24" table. An experimenter sat directly next to them on their left-hand side. The parent sat behind the experimenter and child. A camera next to the table captured the presentation of the stimuli and children's responses.

[3] Analysis including partial data from these 8 participants did not change any of the interpretations of the data discussed below.

Children completed 12 trials. In each trial, children were shown a triad of stimuli (see Figure 9 for example triad). Each triad featured two test characters and a target character. Each of the test characters belonged to one of the same social categories as the target character, but was different from the target character on the other social category. On each trial, the experimenter placed the first test character approximately 10 inches in front of the child. The experimenter then labeled the character's social categories (e.g., "The child is Jewish. He is a boy"). The experimenter then provided a novel fact about the character (e.g., "This Jewish boy likes to play a game called gorp"). This procedure was then repeated with the second test character: the picture was placed 7.5 inches to the right or left side of the first picture (counterbalanced), the experimenter labeled the character's social groups and provided a novel fact about the character (e.g., "This child is Christian. She is a girl. This Christian girl likes to play a game called zool"). Finally, the target character was placed centered approximately 4 inches below the test characters. The experimenter labeled the target character's social categories, noting the similarity with the relevant test character (e.g., "Look at this child. She is Jewish like this child. She is a girl like this child."). The experimenter then asked the child which test character the target character would share characteristics with (the test prompt; e.g., "Does she like to play a game called gorp like him? Or does she like to play a game called zool like her?").

The experimenter pointed to each test character as she was labeling them. When introducing the target character, the experimenter said, "Look at this child". Then, the experimenter labeled the test character's social categories and pointed to the test character that matched the social category being labeled.

If the child did not respond to the test prompt, or pointed to the target character, the experimenter told the child "there are no right or wrong answers, it is okay to guess" up to 2 additional times, pausing approximately 3 seconds after each question. In order to ensure accuracy in recording children's responses, if the child responded verbally the experimenter would ask the child to point to the picture. If the child never responded, the experimenter moved onto the next trial.

Children viewed three types of triads: gender/ethnicity, gender/religion, and religion/ethnicity. We focused on these particular social categories in order to compare the findings with prior studies (Birnbaum et al., 2010; Rhodes & Gelman, 2009; Smyth et al., 2017). The social category labels included in the study were boy/girl, Mexican/White, and Christian/Jewish. We conducted a pilot study to determine the ethnic and religion labels that children in this area might have been exposed to. Twelve parents were asked to indicate whether their 4- to 6-year-old child knew any of the following words: Protestant, Christian, Catholic, Latino, Mexican, Hispanic. Out of the list of words, 50% of children knew the word Christian, 17% of children knew the word Catholic, 0% knew the word Protestant, 58% knew the word Mexican, 17% knew the word Hispanic, and 0% of children knew the word Latino. Because children were unfamiliar with most of the religion terms, we decided to use the word Jewish as a comparison to the religion term Christian, which children were familiar with. Additionally, although Mexican is not an ethnicity, local children were not familiar with Latino or Hispanic, so we used the term that was more familiar to the children.

Novel properties included preferences (e.g., likes to play a game called gorp/zool;), beliefs (e.g., believes that feps/wugs live in trees;), biology (e.g., has blood

type zid/seb), and future occupation aspirations (e.g., wants to be a zagor/dunkel when s/he grows up; see Appendix C for a full list of novel properties). These categories of novel properties were the same as in Birnbaum et al. (2010).

On trials where gender was not relevant (i.e., ethnicity/religion trials), gender was held constant. On two of the ethnicity/religion trials, the stimuli were all boys, and on the other two trials the stimuli were all girls. On trials where ethnicity was not relevant (i.e. gender/religion trials), it was held constant: on two trials the stimuli were all Hispanic, and on the other two trials the stimuli were all White.

Four different randomized orders were created with the constraints that the same triad type was not repeated consecutively, questions based on the same novel property did not occur consecutively, and the test character was not placed on the same side (right or left) more than twice consecutively. For a given trial, the test character that was placed first and the social group label that was provided first were counterbalanced across orders.

Coding. All children were coded offline by a primary and secondary coder, unless parental consent was not given for the session to be recorded. In this circumstance, the experimenter coded the child's responses during the task, and no secondary coding was completed. This only occurred for one participant.

For each trial the coder indicated the child's choice after the test prompt. To ensure accuracy in children's responses, only pointing towards, or physically touching the stimuli were counted as a response. The two coders agreed on all children's test prompt choices. Two measures were calculated: (1) the number of times the child picked a particular social category on each social comparison type (e.g., how many times the child picked ethnicity over gender on the ethnicity vs. gender social comparison) as well as (2) the total number of times the child picked ethnicity, gender, and religion across all social comparison types (total scores).

Results

First, children's performance as a combined group was examined. To investigate children's relative use of social categories when making inductive inferences, chi-square analyses were conducted to assess children's distribution of responses. A separate analysis was conducted for each triad type. For each triad type, we designated one of the two categories as the reference category (i.e. gender for the gender/religion trials) and calculated the number of children who selected that reference category on all four trials, three trials, two trials, one trial, or none of the trials. For instance, we counted the number of children out of the total 64 who on the gender/religion trials picked the gender match on all, three, two, one or none of the trials. The resulting distribution was compared to the distribution that would be expected by chance (four trials: 6.25%; three trials: 25%; two trials: 37.5%; one trial: 25%; and zero trials: 6.25%).

On the gender/religion triads, the distribution of responses differed significantly from chance, χ^2 (4, N = 64) = 11.06, p =.026, with children selecting gender more than religion (see Table 3). Ten children selected gender on all four trials, whereas only two children selected religion on all four trials. On the ethnicity/religion triads, the distribution of responses did not differ from chance, χ^2 (4, N = 64) = 6.20, p = .18, but there was a numeric trend for children to select ethnicity more than religion. On the gender/ethnicity triads, the distribution of responses did not differ from chance, χ^2 (4, N =

64) = 5.75, p = .22, but there was a numeric trend for children to select gender more than ethnicity.

This pattern of findings is broadly consistent with the initial predictions outlined in the introduction. As predicted, children did not demonstrate a clear focus on ethnicity (unlike children in Israel) or religion (unlike children in Northern Ireland) in their choices. The only significant effect that was found was for children to favor gender, as originally predicted. However, this only held true on the gender/religion triads. Surprisingly, children were not significantly more likely to pick gender on the gender/ethnicity triads as originally predicted. One possibility is that this reflects an age effect, with five- and six-year-olds differing in their attention to these social categories. Previous research conducted in the US has shown that in some environments, children's tendency to use gender when making inductive inferences changes with age (e.g., Rhodes & Gelman, 2009). It is therefore possible that children's attention to the categories tested here also differed with age, leading to this mixed pattern of results. To examine this possibility, we investigated whether children's pattern of responses differed across the two age groups.

Children's total scores for each social category (gender, ethnicity, religion) were entered into a multivariate analysis of variance (MANOVA) with age group (5-years-old, 6-year-olds) as a between-subjects factor. The overall MANOVA revealed a marginal effect of age, $F(1, 61) = 2.52$ $p = .089$, $\eta_p^2 = .08$. In particular, age-group had a significant effect on children's selection of ethnicity, $F(1, 62) = 5.07$ $p = .028$, $\eta_p^2 = .08$: 6-year-olds ($M = 4.34$, $SD = 1.26$) were more likely to select ethnicity than were 5-year-olds ($M = 3.53$, $SD = 1.61$).

Given this effect of age, we next examined each age group separately. Chi-square analyses conducted on 5-year-olds distribution of responses on each triad type revealed that children's responses deviated from chance for the two triad types involving gender. In particular, on the gender/religion triads, the distribution of responses differed significantly from chance χ^2 (4, $N = 32$) = 13.68, $p = .008$, with children selecting gender more than religion. Similarly, on the gender/ethnicity triads, the distribution of responses was marginally different from chance χ^2 (4, $N = 32$) = 9.40, $p = .052$, with more children selecting gender than ethnicity. Five-year-olds' distribution of choices did not differ on the religion/ethnicity triads, χ^2 (4, $N = 32$) = .94, $p = .92$, suggesting they picked these categories equally.

In contrast, 6-year-olds distribution of responses only deviated from chance on the religion/ethnicity triads, χ^2 (4, $N = 32$) = 12.02, $p = .017$, with children selecting ethnicity more than religion. Six-year-olds' distribution of responses did not significantly differ from chance on the gender/ethnicity triads, χ^2 (4, $N = 32$) = 4.60, $p = .33$ or the gender/religion triads, χ^2 (4, $N = 32$) = 1.70, $p = .79$, suggesting that children selected each of the social categories equally on these triads.

This pattern of results suggests two age-related changes in children's tendency to use various social categories when making inductive inferences. First, at age 5, children were equally likely to base their inferences on religion and ethnicity. However, at age 6, children were significantly more likely to base their inferences on ethnicity. In combination with the finding that children were more likely to base their inferences on gender than ethnicity at age 5, but not at age 6, these findings suggest that children's

tendency to base their inductive inferences on ethnicity is increasing across the age groups. Second, we found that children were more likely to base their inferences on gender than religion at age 5, but not at age 6. In combination with the finding that children's tendency to base their inferences on ethnicity increases across the age groups, these findings suggests that children's tendency to base their inferences on gender decreases across the age groups.

These results suggest an age-related change in how children reason about social groups. However, an alternative possibility is that these findings reflect other differences between the two groups. For instance, one factor that might influence children's responses is children's knowledge of the various ethnicity and religion terms that were used in the study. An independent samples t-test was conducted to compare the average number of ethnicity words used in the study that 5- and 6-year old children knew, as reported by the parent. The two age groups did not differ on their knowledge of ethnicity words, $t < 1$, or religion words, $t < 1$. Thus, the pattern of results is not likely due to children's differential comprehension of the words across age groups.

Another potential factor that may have influenced the pattern of results was parents' level of education. Prior research has found that children's conceptions of social categories such as gender and race differed across rural, politically conservative communities with lower levels of education and urban, liberal communities with higher levels of education (Rhodes & Gelman, 2009). This raises the possibility that parental education could influence the pattern of results in the current study. To analyze whether parents' level of education differed across the two age groups, parents' level of education was recoded on a four-point scale: high school (1), associates degree (2), bachelors degree (3), and master's or professional degree (4). An independent samples t-test comparing parents' average level of education across age groups revealed there was a trend for the 5-year-old children's parents to be more educated ($M = 2.81$, $SD = 1.12$) than the 6-year-old children's parents ($M = 2.28$, $SD = 1.17$), $t(62) = 1.86$, $p = .07$. We next examined whether this education difference contributed to the differing pattern of responses across age groups. Five- and 6-year-old children's total scores for each social category (gender, ethnicity, religion) were entered into a multivariate analysis of variance (MANOVA) with parent's highest level of education, and age group (5-years-old, 6-years-old) as between-subjects factors. There was a marginal main effect of age, $F(1, 55) = 2.55$ $p = .09$, $\eta_p^2 = .15$. No other effects were significant, both $Fs < 1.18$, $ps > .32$. The absence of any effects or interactions involving parental education suggest that the difference between the two age groups is not a product of the marginal differences in parental education between the two groups. This finding suggests that parents' level of education did not influence the pattern of results seen in Experiment 4[4].

General Discussion

The current study sought to examine how children in a diverse area of the US use religion, ethnicity, and gender when reasoning about the characteristics of others.

[4] An additional analysis found no significant effects of the child's ethnicity (Non-Hispanic White or Other) on children's responses.

Children participated in a forced-choice inference task where they were shown 12 triads of pictures of children; each triad consisted of two test characters from two different social categories (e.g., Jewish boy, Christian girl) and a target character that belonged to one of the same social categories as each test character (e.g., a Christian boy). Children were told novel facts about each of the test characters, and were asked which test character the target character would be most similar to. Children saw three different social comparisons (ethnicity/gender, gender/religion, ethnicity/religion). At age 5, children relied on gender over religion or ethnicity when making inductive inferences about the characteristics of novel individuals. The tendency to use gender was reduced at age 6, where children instead demonstrated an increased tendency to base their inferences on ethnicity. Similar to the 5-year-olds, 6-year-olds were unlikely to make inductive inferences based on religion.

The current study found that at age 5, children tend to use gender as a basis for their inductive inferences over other social categories. The observed pattern of findings is consistent with a number of existing findings that have found that gender is highly salient to children from a young age, and that children tend to focus on this social category when making inductive inferences about the characteristics of others (e.g., Rhodes & Gelman, 2009). In the current study, children's tendency to use gender when making inductive inferences decreased from age 5 to age 6. This age transition is similar to what has been found with children from urban areas of the US (Rhodes & Gelman, 2009; Smyth et al., 2017), where children initially use gender as a basis for inferences, but over time use other social categories. This similarity is interesting, given that the current sample is, overall, more rural, more politically conservative, and less educated than the urban sample used in Rhodes and Gelman (2009). In more rural, conservative areas, children's use of gender when making inductive inferences does not decrease with age (Rhodes & Gelman, 2009). Future work is needed to investigate which aspects of children's environments, and potential local environmental differences, might influence children's attention to gender.

Consistent with our predictions, 5- and 6-year-olds in the current study did not base their inferences on religion. At age 5, children were never more likely to base their inferences on religion over other social categories, demonstrating that religion is not salient to children at this age. Further, the total number of times children picked religion (out of 8 trials) did not significantly change from age 5 (M = 3.63) to age 6 (M = 3.25), suggesting that children as a group are not shifting as group to thinking that religion is highly important for basing their inductive inferences from age 5 to age 6. These findings are similar to what has been found with children growing up in Boston, where children do not use religion significantly more than other social categories with making inductive inferences (Smyth et al., 2017). Yet, this pattern of results is different than what has been found with children in Northern Ireland, where children find religion categories highly salient and use them to make inductive inferences. The difference in findings between children in the US and Northern Ireland could reflect the fact that religion is less salient to children in the US, that areas in the US do not have historical conflicts based on religion and are not typically segregated by religion, or potentially both.

A somewhat complex pattern of findings emerged for children's use of ethnicity when making inductive inferences. At age 5, children did not base their inductive

inferences on ethnicity. This result contrasts with the findings with children from Israel, where children this age tend to base their inferences on ethnicity than other social categories, and supports our prediction that ethnicity is less salient in US children's environments. In countries such as Israel where ethnicity categories are tied to historic, violent conflicts, these categories may become salient to children at a young age. Social categories based on ethnicity may not be as salient to the children from diverse areas in the United States where these categories are not as unavoidably tied to historical conflicts. However, at age 6, children did use ethnicity to make inductive inferences. This increase in children's likelihood to base their inferences on ethnicity is likely what caused the gender/ethnicity comparison to become non-significant at age 6, although there was a trend for 6-year-olds to pick gender (M = 2.19) more than ethnicity (M = 1.81) on the gender/ethnicity trials. Overall, these findings suggest that there is an increase in attention to ethnicity with age.

This pattern of findings was somewhat unexpected, given that previous research with children in Israel and Ireland has found that attending integrated schools with higher levels of social diversity attenuates children's tendency to use certain social categories when making predictions about others (Deeb et al., 2011; Smyth et al., 2017). Consistent with this, frequent positive interactions with out-group members (e.g., members of a different race/ethnicity) that are sustained, frequent, and with a cooperative element, reduces biased attitudes (Allport, 1954; Brown & Hewstone, 2005; Pettigrew & Tropp, 2006). These types of findings in the literature are what led to our prediction that children living in a diverse environment may pay less attention to ethnicity when making inductive inferences. However, it is possible that our sample was less integrated than we thought and that parents and children in the current sample did not have much contact with other ethnic groups. A lot of neighborhoods in the US remain racially segregated (Orfield, 2001), and even in racially diverse areas, children do not always experience friendships with children of different races (Moody, 2001). Perhaps without direct contact and sustained relationships with other ethnic groups, living in a diverse context does not decrease children's tendency to use ethnicity to make inductive inferences.

There is also a conflicting body of findings that find the opposite effect than that of the integrated schools in Israel and Ireland. For instance, White communities with higher levels of racial diversity tend to have higher levels of racial bias (Rae, Newheiser, & Olson, 2015), and immersive exposure to other-race individuals (i.e., African American college students studying in China) does not reduce individuals' explicit racial biases (Qian, Heyman, Quinn, Fu, & Lee, 2017). These findings suggest that mere exposure to diverse contexts may not be sufficient to change children's social-group based attitudes. In fact, it could be that living in a diverse environment may afford additional opportunities for children to notice implicit cues to bias that might happen in cross-group interactions. For instance, Israeli parents that use more generic statements or category labels (e.g., "Arabs are like this") while participating in a storybook task with their 5-year-old, have children that are more likely to endorse essentialist beliefs about ethnicity (Segall, Birnbaum, Deeb, & Diesendruck, 2015). In other words, these children are more likely to form the belief that members of certain ethnic category have inherent, stable properties that make individuals in the same ethnic category similar to one another, but distinct from members of other categories (Medin & Ortony, 1989).

It is also possible that shifts in political discourse relating to ethnicity could have potentially impacted the current results. Children in Experiment 4 were run between April 2017 and March 2019, during a time where ethnicity-related topics were highlghted frequently in the US media and news outlets. This may be a potential subtle environmental influence that lead to children's shift in attention from 5- to 6-years-of age.

In order to fully understand the observed ethnicity pattern in the current study, more information is needed in regards to children's daily experiences with various ethnic groups. Specifically, future work should investigate the level of contact children have with other ethnicities, the nature of this contact, and the level and extent to which children have friendships with children of diverse ethnic backgrounds.

In summary, the findings from the current study highlight the influence of local cultural context on children's social-category based inferences and expand our understanding of children's use of ethnicity when reasoning about social categories. These findings suggest that the relative salience of social categories varies based on environmental context.

Chapter 5

Previous research has found that by preschool, the tendency to assume that members of a social category will share inherent, stable characteristics is well established. Because this tendency is robust by preschool, it suggests that this tendency may emerge earlier. The current dissertation sought to examine the origins of social-group based inferences in infancy, as well as the role of environmental influences in contributing to the development of these inferences across early childhood.

In Project 1, we investigated the nature of infants' social-group based reasoning by examining whether infants generalized behavioral dispositions across members of a social group. The results of Experiment 1 suggested that 20-month-old infants do not yet generalize social dispositions across members of a social group. However, additional results suggested that infants might have difficulty reasoning about social dispositions at this age. Infants in Experiment 1 expected a single individual to be consistent in behaviors when the contexts of the situation remained the same (all her behaviors were in a lego-building context). Yet infants in Experiment 2 did not expect an individual to be consistent in her behaviors when the context of the situation changed (some behaviors were in a lego-building contexts, and some were in a ball-rolling context). Although it is possible that infants might generalize across members of a social group with more support, as discussed in Project 1, the current study provides no evidence that infants generalize behavioral dispositions under the same conditions that infants generalize food preferences at this age (Smith & Scott, 2017).

Prior research has shown that infants attend to characteristics that mark group membership from an early age (e.g., Kelly et al., 2005; Shutts et al., 2009). By the second year of life, infants can use a variety of cues to spontaneously categorize individuals into social groups (Jin & Baillargeon, 2017; Liberman et al., 2016; Powell & Spelke, 2013). Further, some of my own research has shown that at this age, infants generalize some properties, such as food preferences, across members of a social group (Smith & Scott, 2017). Together, these findings suggest that young infants are quite attentive to social groups, and are open to the idea that members of a group may be similar to one another in terms of their psychological properties. Yet, infants' failure to generalize behavioral dispositions in Project 1 suggests that infants do not necessarily know which properties are likely generalizable. This may be something that infants have to learn from their environment.

Children also have to learn which social groups to essentialize from environmental input. The results from Project 1 illustrate that children are not born with essentialist beliefs about social categories. Instead, children begin to form essentialist beliefs about categories in their environments in order to make sense of the complex social world around them. Researchers have proposed that children develop knowledge about the social world in a similar manner to how children develop knowledge about the biological world (Rhodes & Mandalywala, 2017). Children assume that the world is composed of distinct kinds that possess an intrinsic essence that causes members of the same kind to share both inherent, unobservable properties as well extrinsic, observable properties. When reasoning about biological categories, children typically receive consistent information about how to classify different entities (e.g., what makes a dog a different animal than a cat). However, unlike biological categories that are highly stable

over time and contexts (Atran, Estin, Coley, & Medin, 1997), social categories are highly variable and flexible. People can be categorized in many ways. For instance, a person could be categorized by personal characteristics such as race, gender, or social status; by psychological characteristics such as preferences, beliefs, goals, and interests; and even by completely arbitrary characteristics (Rhodes & Mandalywala, 2017). The meanings of these category classifications also vary based on social context and culture. Thus, the boundaries of social categories and the differences between particular categories are much more ambiguous. Because of this ambiguity, children may receive much more inconsistent information about how to structure their social world, and thus turn to influences in the environment to help guide their inferences about social categories.

One type of environmental influence that infants may readily learn about social categories from is how parents talk about social categories. Project 2 investigated whether children could form expectations about a novel social category, Zarpies, after their going through a picture book that either had generic or no-label sentences with their parent. Children that heard the generic statements in the book expected that a novel Zarpie would display similar characteristics to the Zarpie behaviors depicted in the book, whereas children that did not hear the generic statements had no expectation about how a novel Zarpie would behave. In other words, the manner in which parents discussed the social category with their child impacted children's tendency to essentialize the Zarpie category. These findings suggest that parental talk, something that children experience on a daily basis, can have profound effects on the categories children attend to, and the particular beliefs children form about them.

In particular, Project 2 highlights the powerful effect of generic statements as a mechanism by which children learn about social categories. Although parents may not be ostensibly attempting to teach their child something when they use generic statements, children's learning is heavily affected after hearing this type of input. One reason that generic statements powerfully affect children's social-category based inferences and judgments is that generics require very little statistical evidence for acceptance (Abelson & Kanouse, 1966). For instance, it takes very few instances of a shark attacking swimmers for us to make the category-wide generalization that sharks attack people. These category-wide generalizations may be an effective way for children to learn about biological categories in their environment, but when applied to social categories they can have very negative consequences (Leslie, 2017). Further, in Project 2, parents who read the generic statements in the book also made more statements that highlighted how Zarpie members were a distinct group of people than did parents who did not read the generic statements. Thus, hearing the generic statements not only affected the child, but also affected the type of language that parents used when discussing the social category.

These results highlight the importance of generic statements on children's developing beliefs and expectations about social categories. Future research should further investigate potential individual differences in the tendency to produce generic statements. For instance, in Project 2, parents in both conditions produced relatively low instances of generic, comparative, and contrastive statements. However, this was a brief session of parent-child interaction that typically lasted around 5- to 15-minutes. It is possible that with more representative sample, or more instances of how parents discuss

social groups with their children, we would see clearer effects of these types of statements on children's formation of beliefs about social categories. This would allow us to further investigate potential individual and group differences in how parents discuss social categories, as the potential downstream effects on children's learning.

Children must also learn which social categories to attend to when making inductive inferences. As previously mentioned, because the boundaries of social categories are flexible, people can be categorized in many ways. Adding to this complexity, people can simultaneously be members of multiple social categories. For instance, a person can be an athlete and a political activist, or a female and a scientist. Children's learning about which social categories to attend to is likely impacted by children's environment in several ways. For instance, the relative salience of particular social categories varies across cultures. In cultures where social categories are tied to historical conflicts, children may find these categories especially salient, and use them to guide their inferences when make inductive inferences. In Project 3, children in the US were much less likely to attend to religion categories than children in Northern Ireland, where religion categories have been shown to be especially salient in the environment. These results suggest that the relative salience of particular social categories in the environment affects the social categories that children attend to in their environment when making inductive inferences.

Similarly, we found that 5-year-olds in the US are less likely to make inferences based on ethnicity than children in Israel. Somewhat surprisingly, we found that children's tendency to base their inferences on ethnicity increased from age 5 to age 6. Because the children in our study were drawn from a highly diverse area, we expected that children in our study would be less likely to attend to ethnicity based on prior research that found that diverse contexts attenuate children's tendency to use particular social categories (e.g., Deeb et al., 2011; Smyth et al., 2017). Yet, it turns out diversity may have more complex effects than we previously thought. A recent meta-analysis conducted on the relationship between ethnic diversity and social trust found mixed effects: approximately half of the studies reviewed found a significant negative relationship between ethnic diversity and social trust, whereas the other half of the studies reviewed did not (Schaeffer, 2014). This suggests that the relationship between diversity and group-based relationships and reasoning is not straightforward. Conflict theory, proposed by Allport (1954), predicts that mere exposure to groups with different ethnic backgrounds can cause out-group prejudice and in-group favoritism (Bobo & Hutchins, 1996; Quillan, 1995). If children do not have sustained, meaningful interactions with individuals of other ethnicities, it is possible that ethnicity becomes a salient dimension to children. In fact, several studies found that when individuals have meaningful contact with diverse ethnic groups, the impact of diversity on social trust becomes less negative (Dinesen and Sønderskov 2015; Stolle, Soroka, & Johnston, 2008). Further, some have argued that people may live in a broader diverse context (e.g., California) while simultaneously live in a rather homogenous local environmental context (Uslaner, 2012). In these contexts, there may not be as many opportunities to interact with individuals of other backgrounds, which has been found to reduce biased attitudes (Allport, 1954; Brown & Hewstone, 2005; Pettigrew & Tropp, 2006). Overall, these findings suggest the impact of diversity on children's tendency to use ethnic categories when reasoning about

others is highly complex. More research is needed on the amount of contact children have with individuals of other ethnicities, as well as the nature of these interactions in order to clarify this complex relationship between diversity and children's tendency to use categories based on ethnicity when reasoning about others.

Although children may readily attend to social groups from early on in life, input from the environment influences the types of characteristics children generalize across social groups and the particular social groups children attend to in the environment. However, the biases that children form about social groups are not inevitable, as shown by Project 2 and in Project 3. In Project 2, children that did not hear the generic statements did not believe that Zarpies members would behave similarly. This provides us with a potential avenue for diminishing the potential negative downstream effects of essentialist thinking by changing the way we talk about social categories. Project 3 demonstrated that children do not have strong beliefs about religion at either age point we tested. These findings suggest that children are not born with biases about social categories. Children must be learning these biases, and they only learn them under particular kinds of environmental conditions. Thus, forming categories and using them to inform our inferences does not always have negative effects. On the contrary, categorization is a fundamental piece of human cognition that is extremely important and valuable. To summarize, it is clear that categorization does not inevitably lead to these biases. Biases happen under particular conditions, and the studies investigated here provide a first step for knowing what these conditions are.

References

Abelson, R. P., & Kanouse, D. E. (1966). Subjective acceptance of verbal generalizations. In S. Feldman (Ed.), *Cognitive consistency: Motivational antecedents and behavioral consequents* (pp. 171-197). New York: Academic Press.

Agerström, J., Björklund, F., Carlsson, R., & Rooth, D. O. (2012). Warm and competent Hassan= cold and incompetent Eric: A harsh equation of real-life hiring discrimination. *Basic and Applied Social Psychology*, *34*, 359-366.

Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.

Atran, S., Estin, P., Coley, J. D., & Medin, D. L. (1997). Generic species and basic levels: Essence and appearance in folk biology. *Journal of Ethnobiology*, *17*, 22-45.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412.

Bar-Tal, D., & Teichman, Y. (2005). *Stereotypes and prejudice in conflict: Representations of Arabs in Israeli Jewish society*. New York, NY, US: Cambridge University Press.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1-48.

Bigler, R. S., Jones, L. C., & Lobliner, D. B. (1997). Social categorization and the formation of intergroup attitudes in children. *Child Development*, *68*, 530-543.

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

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

Bobo, L., & Hutchings, V. L. (1996). Perceptions of racial group competition: Extending Blumer's theory of group position to a multiracial social context. *American Sociological Review*, *61*, 951–972.

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

Brown, R., & Hewstone, M. (2005). An integrative theory of intergroup contact. *Advances in Experimental Social Psychology*, *37*, 255-343.

Buttelmann, D., Zmyj, N., Daum, M., & Carpenter, M. (2013). Selective imitation of in‐group over out‐group members in 14‐month‐old infants. *Child Development*, *84*, 422-428.

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

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

Dinesen, P. T., & Sønderskov, K. M. (2015). Ethnic diversity and social trust evidence from the micro-context. *American Sociological Review, 80*, 550–573.

Duh, S., Wang, S., & Goldman, E. J. (2017, April). Stronger intentional cues facilitate infants'attribution of social dispositions, in the symposium *Developmental progression of children's reasoning about others' dispositions and intentions* at the biennial meeting of the Society for Research in Child Development, Austin, TX.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences*, *99*, 15822-15826.

Gallagher, T. (2010). Building a shared future from a divided past: Promoting peace through education in Northern Ireland. In G. Salomon & E. Cairns, (Eds.), *Handbook of peace education* (pp. 241–251). New York, NY: Psychology Press.

Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology*, 20, 65-95.

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

Gelman, S. A. (2004). Learning words for kinds: Generic noun phrases in acquisition. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 445–483). Cambridge, MA: MIT Press.

Gelman, S. A., Coley, J. D., Rosengren, K., Hartman, E., & Pappas, A. (1998). Beyond labeling: The role of maternal input in the acquisition of richly-structured categories. *Monographs of the Society for Research in Child Development*, *63*(1, Serial No. 253).

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., Goetz, P. J., Sarnecka, B. S., & Flukes, J. (2008). Generic language in parent-child conversations. *Language Learning and Development, 4*, 1-31.

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

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*, *58*, 1532-1541.

Gelman, S. A., & Raman, L. (2003). Preschool children use linguistic form class and pragmatic cues to interpret generics. *Child Development, 24*, 308-325.

Gelman, S. A., Ware, E. A., Kleinberg, F., Manczak, E. M., & Stilwell, S. M. (2014). Individual differences in children's and parents' generic language. *Child Development, 85*, 924-940.

Gillespie, G. (2010). *A short history of the troubles*. Dublin, Ireland: Gill & Macmillan.

Graham, S. A., Kilbreath, C. S., & Welder, A. N. (2001). Words and shape similarity guide 13-month-olds' inferences about nonobvious object properties. In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 352-357).

Hamlin, J. K. (2013). Failed attempts to help and harm: Intention versus outcome in preverbal infants' social evaluations. *Cognition*, *128*, 451-474.

Hamlin, J. K. & Wynn, K. (2011). Young infants prefer prosocial to antisocial others. *Cognitive Development*, *26*, 30-39.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*, 557-559.

Hamlin, J. K., Wynn, K., Bloom, P. (2010). 3-month-olds show a negativity bias in social evaluation. *Developmental Science*, *13*, 923-939.

Hamlin, J. K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Sciences*, *108*, 19931-19936.

Haslam, N., Rothschild, L., & Ernst, D. (2002). Are essentialist beliefs associated with prejudice? *British Journal of Social Psychology*, *41*, 87-100.

Heyman, G. D., & Gelman, S. A. (1998). Young children use motive information to make trait inferences. *Developmental Psychology*, *34*, 310-321.

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

Howard, L. H., Henderson, A. M., Carrazza, C., & Woodward, A. L. (2015). Infants' and young children's imitation of linguistic in-group and out-group informants. *Child Development*, *86*, 259-275.

Jin, K. S., & Baillargeon, R. (2017). Infants possess an abstract expectation of ingroup support. *Proceedings of the National Academy of Sciences*, *114*, 8199-8204.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Gibson, A., Smith, M., ... & Pascalis, O. (2005). Three‐month‐olds, but not newborns, prefer own‐race faces. *Developmental Science*, *8*, F31-F36.

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

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

Koch, B., Kenward, B., Fawcett, C., & Gredebäck, G. (2015). *Introducing live eye-tracking with motion-capture for automatically tracked areas of interest (AOI): An application in the study of infant social cognition.* Poster presented at the International Symposium on Action Research, Nijmegen, Netherlands.

Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science*, *14*, 402-408.

Leslie, S. J. (2017). The original sin of cognition: Fear, prejudice and generalization. *The Journal of Philosophy, 114,* 1-29.

Leslie, S. J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, *347*, 262-265.

Liberman, Z., Woodward, A. L., Sullivan, K. R., & Kinzler, K. D. (2016). Early emerging system for reasoning about the social nature of food. *Proceedings of the National Academy of Sciences*, *113*, 9480-9485.

Lloyd, C. D., & Shuttleworth, I. G. (2012). Residential segregation in Northern Ireland in 2001: Assessing the value of exploring spatial variations. *Environment & Planning A*, *44*, 52–67.

Lloyd, K., & Robinson, G. (2011). Intimate mixing—Bridging the gap? Catholic-Protestant relationships in Northern Ireland. *Ethnic and Racial Studies*, *34*, 2134–2152.

LoBue, V., & Thrasher, C. (2015). The child affective facial expression (CAFE) set: Validity and reliability from untrained adults. *Emotion Science*, *5*, 1532.

Mahajan, N., & Wynn, K. (2012). Origins of "us" versus "them": Prelinguistic infants prefer similar others. *Cognition*, *124*, 227-233.

Martin, C. L., Eisenbud, L., & Rose, H. (1995). Children's gender‐based reasoning about toys. *Child Development*, *66*, 1453-1471.

Medin, D., Ojalehto, b., Waxman, S., & Bang, M. (2015). Relations: Language, epistemologies, categories and concepts. In E. Margolis & S. Laurence (Eds.), *The conceptual mind: New directions in the study of concepts* (pp. 349-378). Cambridge: MIT Press.

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

Moody, J. (2001). Race, school integration, and friendship segregation in America. *American Journal of Sociology*, *107*, 679–716.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*, 16474-16479.

Orfield, G. (2001). *Schools more separate: Consequences of a decade of resegregation*. Cambridge, MA: The Civil Rights Project.

Pappas, A., & Gelman, S. A. (1998). Generic noun phrases in mother–child conversations. *Journal of Child Language*, *25*, 19-33.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1), 8-13.

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8‐month‐old infants. *Child Development*, *80*, 674-685.

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*, 751-783.

Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. *Proceedings of the National Academy of Sciences*, *110*, E3965-E3972.

Premack, D., & Premack, A. J. (1997). Infants attribute value ± to the goal-directed actions of self-propelled objects. *Journal of Cognitive Neuroscience*, *9*, 848-856.

Prentice, D. A., & Miller, D. T. (2007). Psychological essentialism of human categories. *Current Directions in Psychological Science*, *16*, 202-206.

Qian, M. K., Quinn, P. C., Heyman, G. D., Pascalis, O., Fu, G., & Lee, K. (2017). Perceptual individuation training (but not mere exposure) reduces implicit racial bias in preschool children. *Developmental Psychology*, *53*, 845-859.

Quillian, L. (1995). Prejudice as a response to perceived group threat: Population composition and anti-immigrant and racial prejudice in Europe. *American Sociological Review*, *60,* 586–611.

R Core Team. (2014*). R: A language* and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rae, J. R., Newheiser, A. K., & Olson, K. R. (2015). Exposure to racial out-groups and implicit race bias in the United States. *Social Psychological and Personality Science*, *6*, 535-543.

Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology*, *59*, 244-274.

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

Rhodes, M., & Mandalaywala, T. M. (2017). The development and developmental consequences of social essentialism. *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*, e1437.

Richman, C. L., Bovelsky, S., Kroovand, N., Vacca, J., & West, T. (1997). Racism 102: The classroom. *Journal of Black Psychology*, *23*, 378-387.

Roberts, S. O., & Gelman, S. A. (2015). Do children see in Black and White? Children's and adults' categorizations of multiracial individuals. *Child Development*, *86*, 1830-1847.

Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kinds? In G.R. Semin & K. Fiedler (Eds.), *Language, Interaction and Social Cognition* (pp. 11-36). London: Sage.

Schaeffer, M. (2014). *Ethnic diversity and social cohesion: Immigration, ethnic fractionalization and potentials for civic action*. Aldershot, UK: Ashgate.

Segall, G., Birnbaum, D., Deeb, I., & Diesendruck, G. (2015). The intergenerational transmission of ethnic essentialism: How parents talk counts the most. *Developmental Science*, *18*, 543-555.

Shutts, K., Kinzler, K. D., & DeJesus, J. M. (2013). Understanding infants' and children's social learning about foods: Previous research and new prospects. *Developmental psychology*, *49*, 419-425.

Shutts, K., Kinzler, K. D., McKee, C. B., & Spelke, E. S. (2009). Social information guides infants' selection of foods. *Journal of Cognition and Development*, *10*, 1-17.

Shutts, K., Roben, C. K. P., & Spelke, E. S. (2013). Children's use of social categories in thinking about people and social relationships. *Journal of Cognition and Development*, *14*, 35-62.

Smith, M. A., & Scott, R. M. (2017). 20-month-olds use social-group membership to make inductive inferences. *Proceedings of the 39th Annual Conference of the Cognitive Science Society*. London, UK: Cognitive Science Society.

Smyth, K., Feeney, A., Eidson, R. C., & Coley, J. D. (2017). Development of essentialist thinking about religion categories in Northern Ireland (and the United States). *Developmental Psychology*, *53*, 475–496.

Stolle, D., Soroka, S., & Johnston, R. (2008). When does diversity erode trust? Neighborhood diversity, interpersonal trust and the mediating effect of social interactions. *Political Studies, 56,* 57–75.

Storage, D., Horne, Z., Cimpian, A., & Leslie, S. J. (2016). The frequency of "Brilliant" and "Genius" in teaching evaluations predicts the representation of women and African Americans across fields. *PloS one*, *11*, e0150194.

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

Taylor, M. G., Rhodes, M., & Gelman, S. A. (2009). Boys will be boys; cows will be cows: Children's essentialist reasoning about gender categories and animal species. *Child Development*, *80*, 461-481.

Tindall, T., & Hamil, B. (2004). Gender disparity in science education: The causes, consequences, and solutions. *Education*, *125*, 282-296.

Uslaner, E. M. (2012). *Segregation and mistrust: Diversity, isolation, and social Cohesion*. New York: Cambridge University Press.

Waxman, S. R. (2013). Building a Better Bridge. In M. Banaji, S. Gelman & S. Lehr (Eds.), *Navigating the social world: The early years* (pp. 292-296). New York, NY: Oxford University Press.

Appendix A: Essentialist Belief-inducing Paragraph Used in Experiment 3

"Imagine that some scientists went to a far- away place and they discovered some people living there. These were people that the scientists had never studied before, so the scientists studied the people in detail. They looked at all of their biological features—their DNA and their blood types—and at their cultural practices. They discovered that the population was very different from any community that scientists had ever studied before. They had very different/ biological properties, and very different cultural practices, from any people that the scientists had ever studied. The scientists had discovered Zarpies!"

Appendix B: Vocabulary Checklist Used in Experiment 4

Child's Name: _____

Child's Birthdate: _____

Parent filling out this form: _____

Please indicate whether your child understands the following words:

Protestan	☐	Yes	☐	No
Christian	☐	Yes	☐	No
Catholic	☐	Yes	☐	No
Jewish	☐	Yes	☐	No
White	☐	Yes	☐	No
Latino	☐	Yes	☐	No
Mexican	☐	Yes	☐	No
Hispanic	☐	Yes	☐	No

Appendix C: Novel Properties Used in Experiment 4

Category	Properties
Preferences	Likes to play a game called gorp/zool
	Likes to pim/nade with his/her friend
	Likes to read books about neem/glip
Beliefs	Believes that feps/wugs live in trees
	Thinks it is important to be framish/blickish
	Thinks that daxes eat quibs/rabs
Biology	Has blood type zid/seb
	Has lindon/kiron in his/her stomach
	Has zorbs/flarns inside of him/her
Future occupation aspirations	Wants to be a zagor/dunkel when s/he grows up
	Wants to be ludine/zaber s/he grows up
	Plans to be a jeckiff/naggle in the future

Table 1

Text Used in the Picture-book Task (Experiment 3)

Page	Generic	No label
1	Look at this Zarpie! Zarpies love to eat flowers	Look at this! This one loves to eat flowers.
2	Look at this Zarpie! Zarpies have stripes in their hair.	Look at this! This one has stripes in her hair.
3	Look at this Zarpie! Zarpies can bounce a ball on their heads.	Look at this! This one can bounce a ball on his head.
4	Look at this Zarpie! Zarpies like to sing.	Look at this! This one likes to sing.
5	Look at this Zarpie! Zarpies climb tall fences.	Look at this! This one climbs tall fences.
6	Look at this Zarpie! Zarpies flap their arms when they are happy.	Look at this! This one flaps her arms when she is happy.
7	Look at this Zarpie! Zarpies have freckles on their feet.	Look at this! This one has freckles on his feet.
8	Look at this Zarpie! Zarpies hop over puddles.	Look at this! This one hops over puddles.
9	Look at this Zarpie! Zarpies hate walking in the mud.	Look at this! This one hates walking in the mud.
10	Look at this Zarpie! Zarpies draw stars on their knees.	Look at this! This one draws stars on her knees.
11	Look at this Zarpie! Zarpies can flip in the air.	Look at this! This one can flip in the air.
12	Look at this Zarpie! Zarpies are scared of ladybugs.	Look at this! This one is scared of ladybugs.
13	Look at this Zarpie! Zarpies hate ice cream.	Look at this! This one hates ice cream.
14	Look at this Zarpie! Zarpies chase shadows.	Look at this! This one chases shadows.
15	Look at this Zarpie! Zarpie babies are wrapped in orange blankets.	Look at this! This one baby is wrapped in an orange blanket.
16	Look at this Zarpie! Zarpies sleep in tall trees.	Look at this! This one sleeps in tall trees.

Table 2

Descriptive Statistics for Parental Talk in Experiment 3, Separately by Picture-book Condition in Experiment 3

	Generic Condition	No-label Condition
	Mean (SD)	Mean (SD)
Percentage of Statements		
Ball talk	6.50 (2.16)	5.94 (1.42)
Ice cream talk	7.24 (2.50)	7.20 (1.71)
Flap/spin talk	6.16 (2.17)	5.98 (1.79)
Generic statements	1.49 (1.89)	0.24 (1.02)
Comparative statements	1.90 (1.90)	0.85 (1.26)
Contrastive statements	2.36 (3.23)	1.15 (1.81)
Number of Statements		
Ball talk	7.47 (4.70)	5.97 (2.77)
Ice cream talk	7.81 (3.64)	7.42 (3.97)
Flap/spin talk	6.44 (2.51)	6.10 (3.48)
Generic statements	1.66 (2.19)	0.19 (0.65)
Comparative statements	2.03 (1.93)	0.87 (1.36)
Contrastive statements	2.78 (3.92)	1.03 (1.64)
Total number of utterances	110.22 (41.18)	101.87 (46.37)

Table 3

The Number of Children in Experiment 4 Who Selected a Particular Social Category on All Four Trials, Three Trials, Two Trials, One Trial, or None of the Trials, Separately by Triad Type and Age Group.

Gender vs. Ethnicity Trials

	5-year-olds		6-year-olds		Overall	
	Gender	Ethnicity	Gender	Ethnicity	Gender	Ethnicity
Number of selections						
0	1	6	2	1	3	7
1	7	9	4	12	11	21
2	9	9	13	13	22	22
3	9	7	12	4	21	11
4	6	1	1	2	7	3

Gender vs. Religion Trials

	5-year-olds		6-year-olds		Overall	
	Gender	Religion	Gender	Religion	Gender	Religion
Number of selections						
0	1	7	1	3	2	10
1	6	8	7	10	13	18
2	10	10	11	11	21	21
3	8	6	10	7	18	13
4	7	1	3	1	10	2

Ethnicity vs. Religion Trials

	5-year-olds		6-year-olds		Overall	
	Ethnicity	Religion	Ethnicity	Religion	Ethnicity	Religion
Number of selections						
0	3	2	1	6	4	8
1	8	9	3	10	11	19
2	10	10	12	12	22	22
3	9	8	10	3	19	11
4	2	3	6	1	8	4

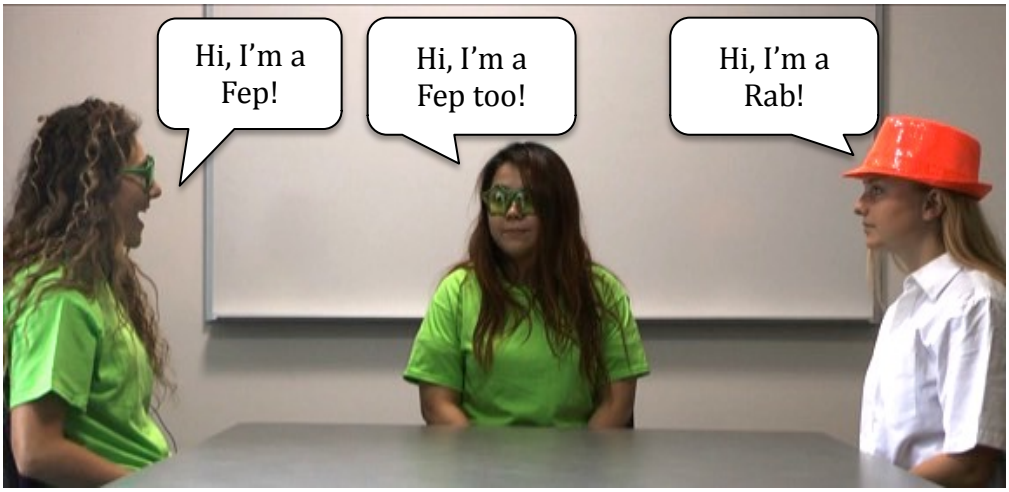
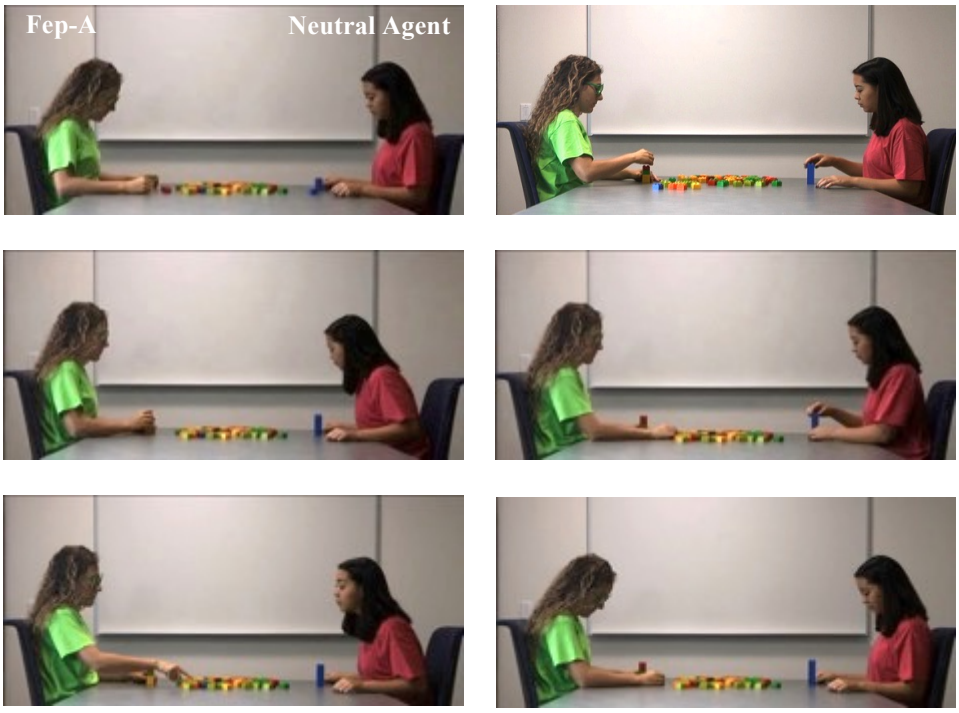


Figure 1. Group-induction trial from Experiment 1.

Prosocial Familiarization Trial 1



Antisocial Familiarization Trial 1

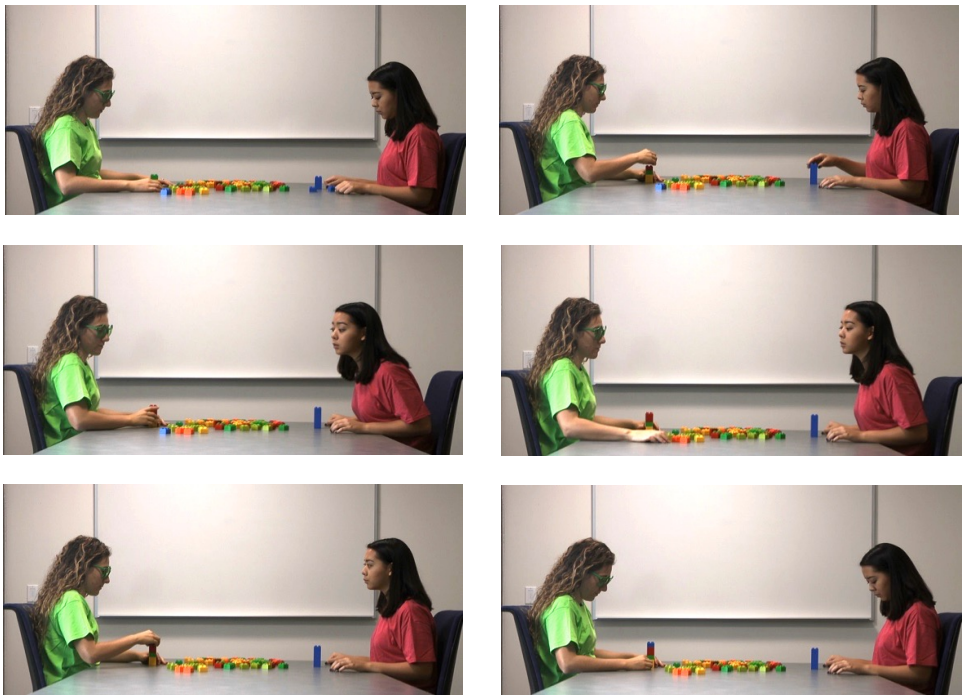
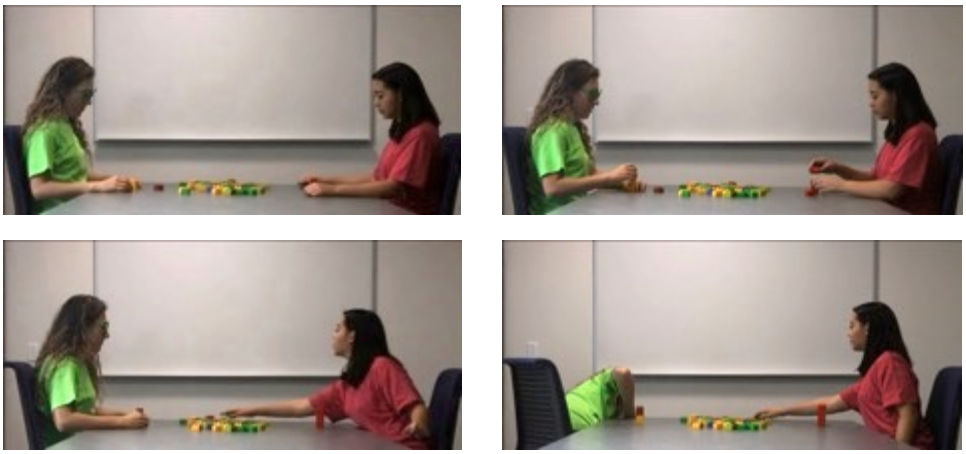
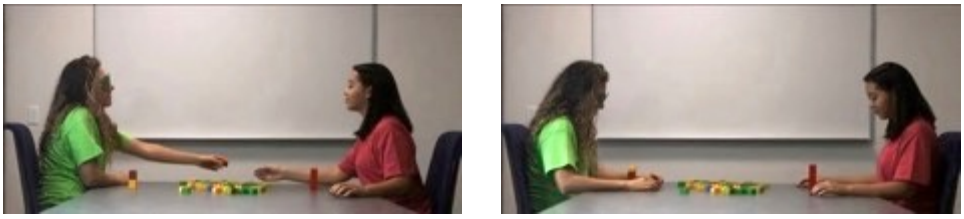


Figure 2. Sample prosocial and antisocial familiarization trials from Experiment 1. Infants viewed either three prosocial familiarization trials or three antisocial familiarization trials.

Test Trial



Prosocial Event



Antisocial Event

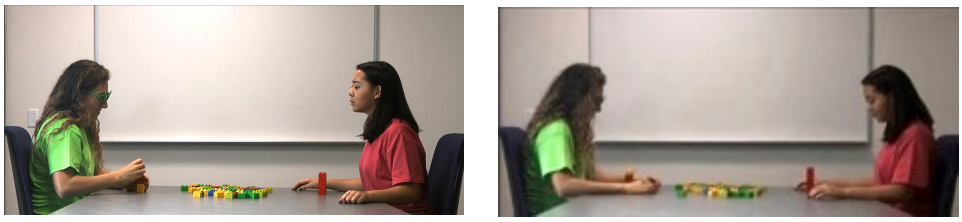


Figure 3. Prosocial and antisocial test events from same-agent condition of Experiment 1.

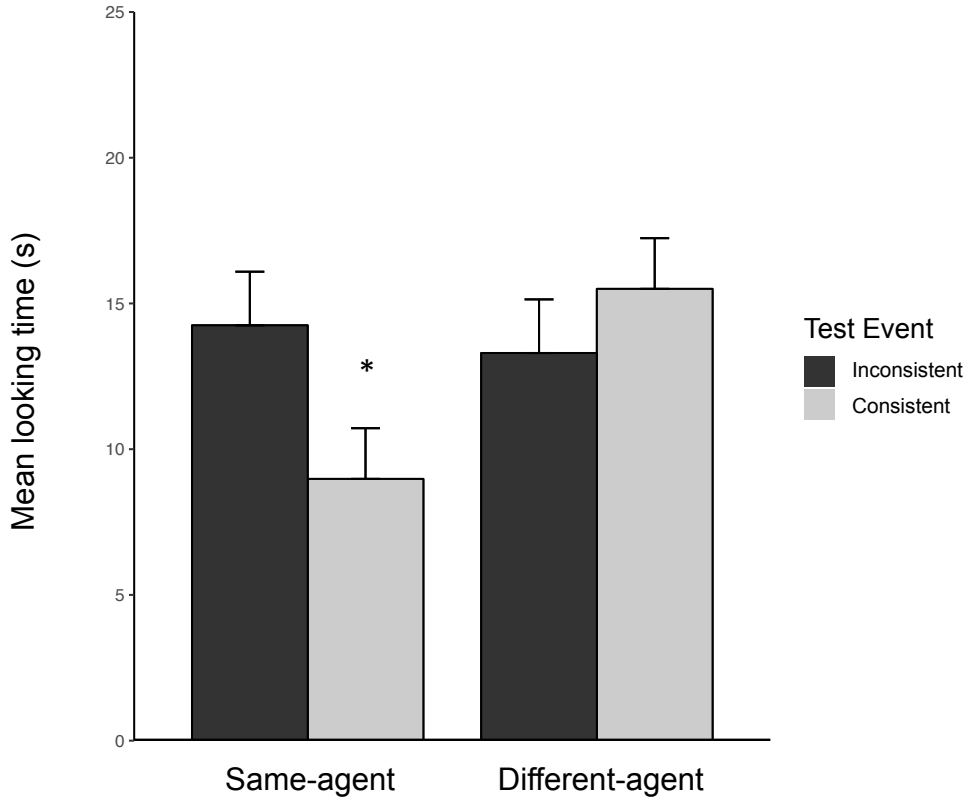
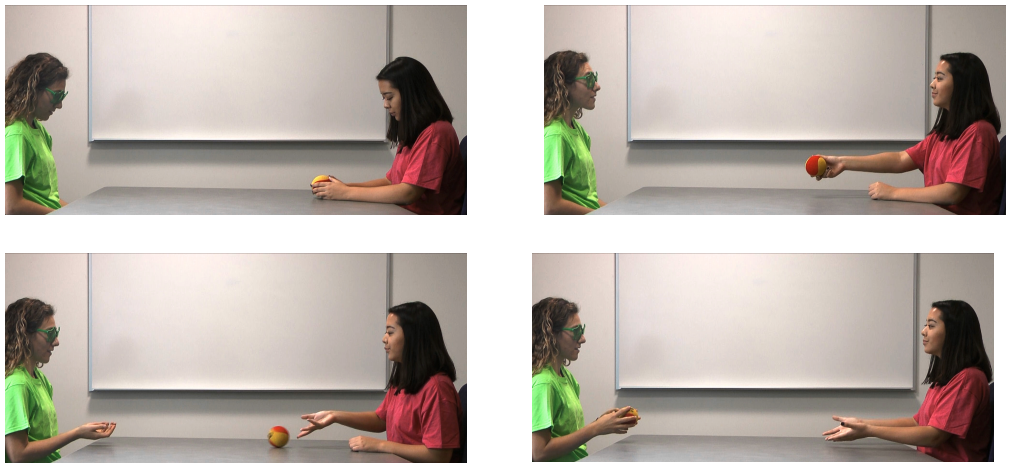
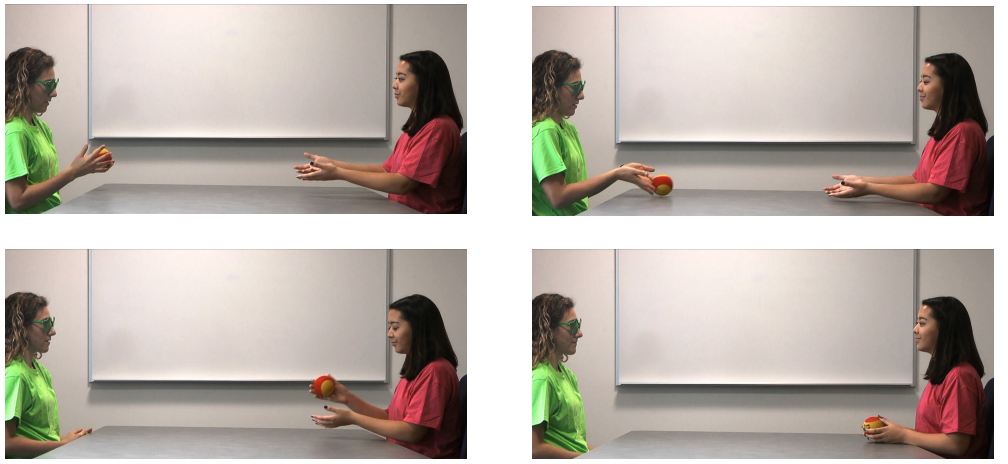


Figure 4. Infants' mean looking time (in seconds) to the inconsistent and consistent test events in the same-agent and different-agent conditions of Experiment 1.

Test Trial



Prosocial Event



Antisocial Event

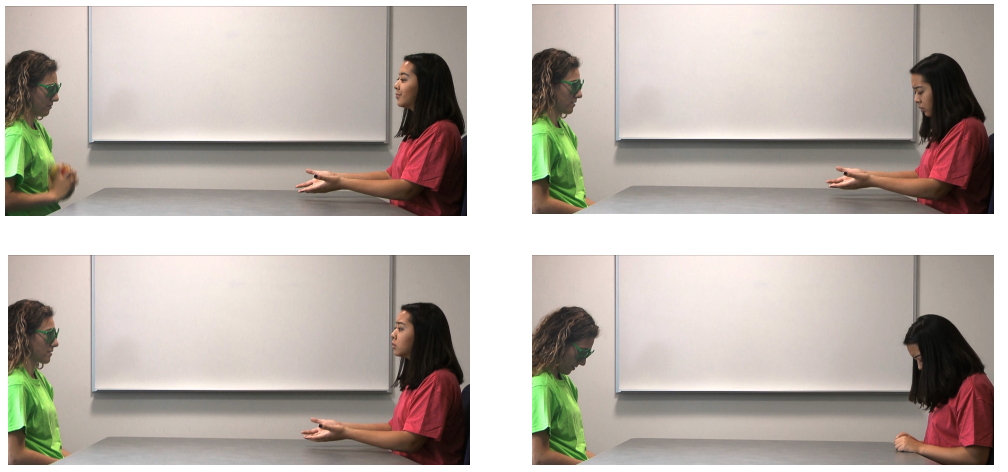


Figure 5. Prosocial and antisocial test events from Experiment 2.

Page 3	Look at this Zarpie! Zarpies can bounce a ball on their heads.	
Page 6	Look at this Zarpie! Zarpies hate ice cream.	
Page 13	Look at this Zarpie! Zarpies flap their arms when they are happy.	

Figure 6. Sample pages from the picture-book task in Experiment 3.

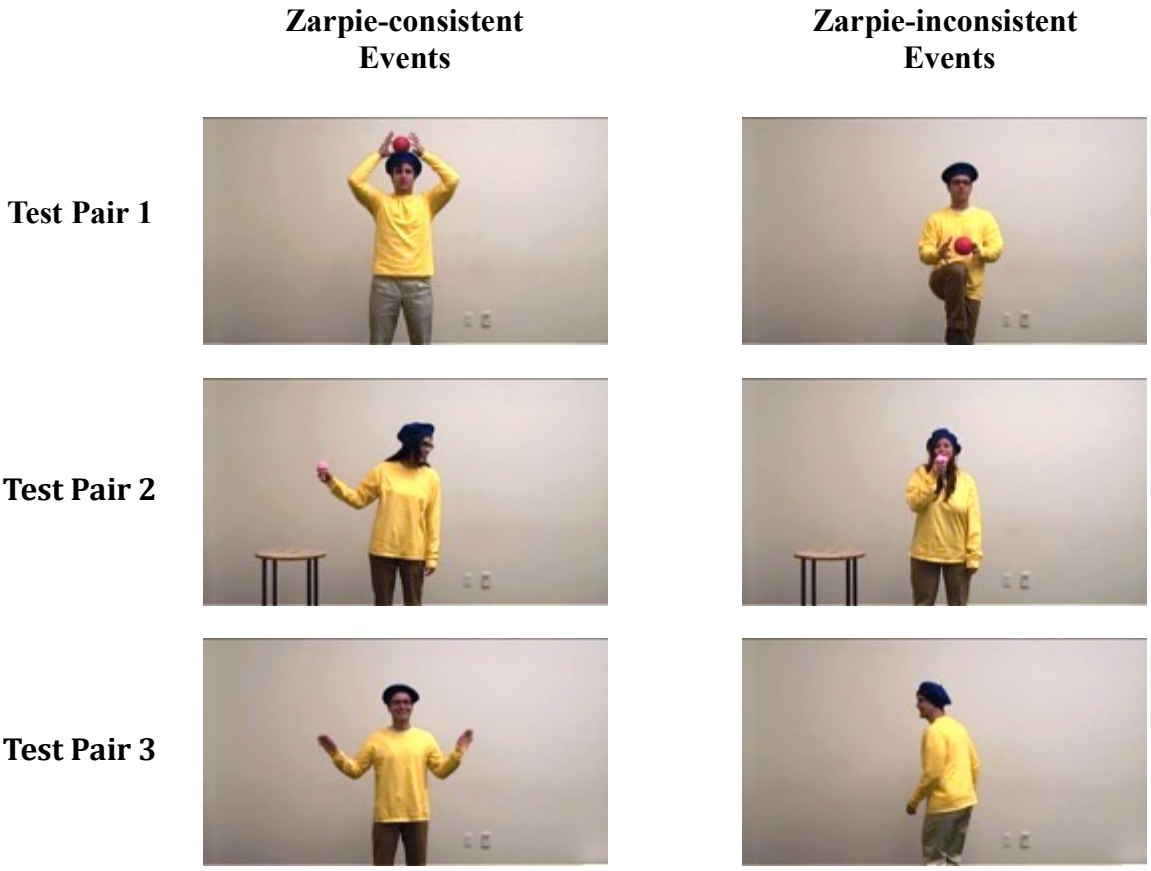


Figure 7. Test events shown in the video task of Experiment 3. Children saw pairs of events in which a person dressed as a Zarpie displayed behaviors that were either consistent, or inconsistent with the Zarpie behaviors depicted in the book (order counterbalanced).

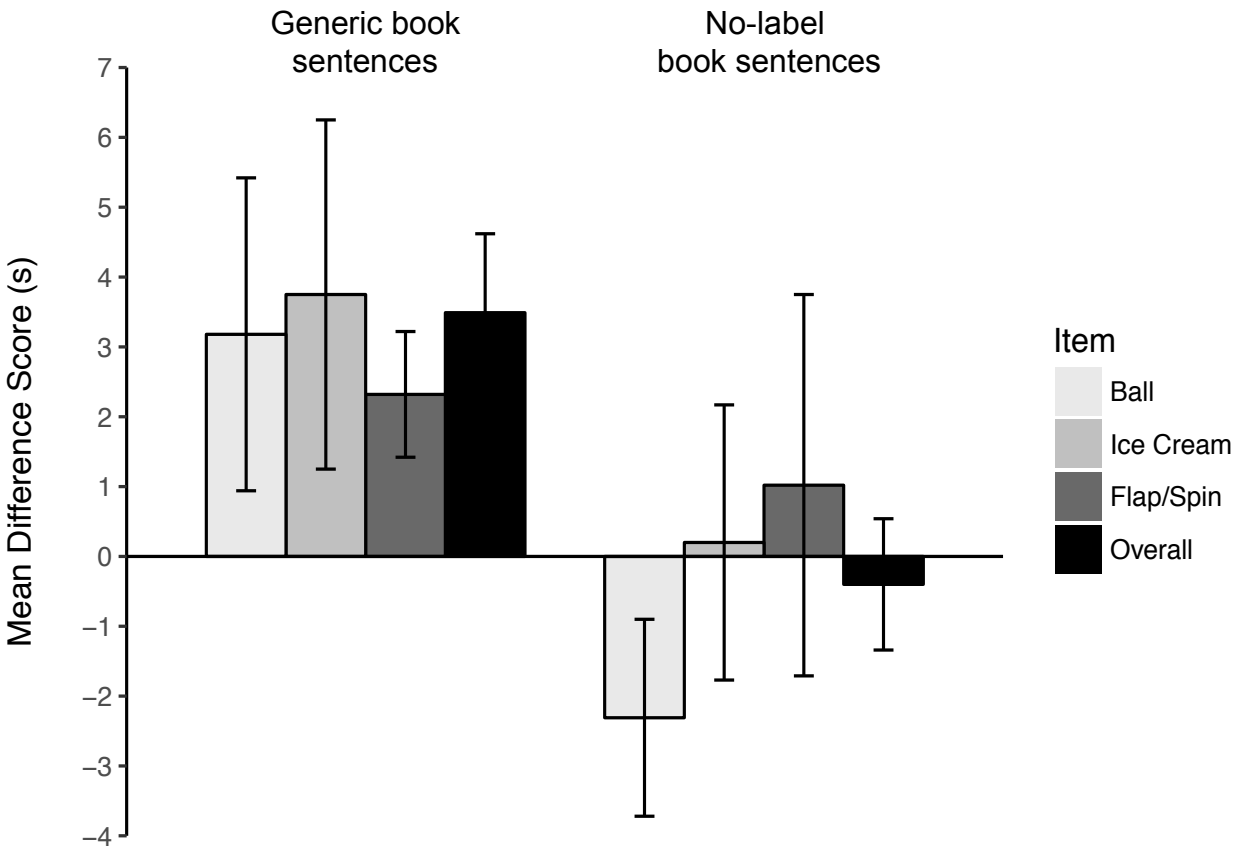


Figure 8. Children's difference scores for the ball, ice cream, and flap/spin items in the video task, as well as children's overall difference score for both the generic and no-label picture-book sentences in Experiment 3.

This child is Jewish. He is a boy. This Jewish boy likes to play a game called Gorp.

This child is Christian. She is a girl. This Christian girl likes to play game called Zool.

Look at this child. She is Jewish like this child. She is a girl like this child. Does she like to play a game called Gorp like him? Or does she like to play a game called Zool like her?

Figure 9. Example of a religion/gender trial from Experiment 4.