SOCIO-MORAL EXPECTATIONS IN INFANTS AND TODDLERS

BY

ZIJING HE

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Professor Renée Baillargeon, Chair
Professor Dov Cohen
Professor Cynthia Fisher
Professor Ying-Yi Hong, Nanyang Technological University
Professor Peggy Miller
Associate Professor Glenn Roisman
ABSTRACT

The present research explored whether infants and toddlers would have early expectations of socio-moral norms. In particular, we asked whether and how children’s expectations about reciprocity would be modulated by considerations of ingroup loyalty. Chapter 2 focused on the idea of “escalation” and provided evidence that infants expect an individual to retaliate more severely against an outgroup member than against an ingroup member. Chapter 3 focused on the idea of “co-retaliation” and suggested that both infants and toddlers expect an individual to retaliate against an outgroup member who had previously hindered the individual’s group member. Chapter 4 focused on the idea of “privilege” and showed that toddlers expect an individual to act more positively in response to a friendly overtue by an ingroup member than by an outgroup member.

Experiments described in this dissertation provide converging evidence that infants and toddlers have rudimentary socio-moral expectations about actions and interactions within social contexts; they are sensitive to various markers of social groups; and their expectations of reciprocity are modulated by considerations of ingroup loyalty. These socio-moral expectations emerge early in life and are likely to be based on a small set of innate socio-moral principles.
ACKNOWLEDGEMENTS

This project would not have been possible without the support and help of many people. Many thanks to my dear advisor, Renée Baillargeon, who offered me tremendous help with my research. Also thanks to my other committee members, Dov Cohen, Cynthia Fisher, Ying Yi Hong, Peggy Miller, and Glenn Roisman, who generously offered guidance and support. Thanks to the grant from NICHD (HD-021104) and the dissertation completion fellowship from the graduate college of the University of Illinois at Urbana-Champaign, providing the financial means to complete this project. Many thanks to Jaclyn Aldridge, Amélie Bernard, Danielle Parrish, Laura Kerlin, Audrey Kittredge, Rose Scott, Pei Pei Setoh, Stephanie Sloane, Maayan Stavans, Di Wu, Melissa Yako, and Sylvia Yuan, who help me collect data and commented on this project. Thanks to all the undergraduates who worked in the Infant Cognition Laboratory and thanks to all the families who had their children participate in the studies. Finally, thanks to my parents, my love Yan, and numerous friends who always offer support and love.
# TABLE OF CONTENTS

- **CHAPTER 1: INTRODUCTION** .......................................................... 1
- **CHAPTER 2: PROJECT 1 ON ESCALATION** ........................................... 20
- **CHAPTER 3: PROJECT 2 ON CO-RETAIATION** ................................... 36
- **CHAPTER 4: PROJECT 3 ON PRIVILEGE** .......................................... 64
- **CHAPTER 5: GENERAL DISCUSSION** .............................................. 78
- **FIGURES** ......................................................................................... 89
- **REFERENCES** .................................................................................. 116
CHAPTER 1: INTRODUCTION

When young children are in social situations, such as on the playground, their parents often remind them: “wait for your turn,” “don’t push,” “say thank you,” “be nice to your friends,” and so on. Also, in children's literature, as in Aesop's fable or Grimm's Fairy Tales for example, a moral lesson is often included in order to teach children social and moral norms. But when do children start to show sensitivity to norms? And what norms do young children first apply to interactions among individuals?

1. Views on the origin of morality

The origin of morality has been one of the most hotly debated topics in intellectual history. Psychologists have formed two major views about where our moral sense comes from. The late-acquisition view states that the acquisition of morality does not begin until the preschool years or even later; and hence children come into the world with a very limited innate basis for socio-moral understanding (note: psychologists coined the term socio-morality to refer to moral behaviors and judgments that occur within a social context). In contrast, the early-emergence view states that children are born with a rich innate basis for understanding social interactions and hence sensitivity to moral norms emerges early in life. Below is a brief review of both the late-acquisition and the early-emergence views.

1.1. Late-acquisition view

The late-acquisition view argues that children are not born with a moral sense; instead, they acquire morality through experience. This view has been studied in two main approaches. The socialization-based approach examines how parental discipline practices and other socialization processes help children internalize and conform to societal norms. In contrast, the construction-based approach adopts a more cognitive stance that explores how children gradually construct moral concepts and norms through social interactions.

1.1.1. Socialization-based approach

The socialization-based approach focuses on how social experiences make children behave in ways that are consistent with societal standards, or even selfless. According to the psychoanalytic theory,
great tension exists between an individual and society. For Freud (1923/1962), the tension is rooted in the incompatibility of the individual’s instinctual impulses and society’s function of protecting its members from individual aggressive tendencies. People develop a moral sense when their instinctual need for self-gratification becomes transformed through the acquisition of a conscience. Children identify with the parent of the same sex and internalize parental and cultural values. These values are incorporated into children’s conscience and become their moral guide. Children also internalize guilt, shame, fear and anxiety as the means for regulating behaviors.

Learning theory conceptualizes the acquisition of morality in a different way. Skinner (1971) argued that people develop a moral sense as their behaviors are reinforced or punished by value judgments that are based on societal standards. Behaviors themselves are not intrinsically good or bad. Children learn to identify acts as legal, illegal, pious, or sinful as a consequence of contingencies of reinforcement. Learned behaviors do not reflect the nature of a person’s character, but are constrained by effective contingencies in the environment.

Both the psychoanalytic and learning theories focus on the central role of aversive emotions (e.g., fear, anxiety, shame, and guilty) in the acquisition of morality. As the influences of these theories waned over the years, the scope of inquiry was broadened to include and even emphasize the role of the positive emotions (e.g., attachment, bonding, love, sympathy and empathy) in socio-moral development. For example, research has found that maternal responsiveness (contingent, appropriate responding) to their infants at 9 months predicted higher levels of empathic responses in toddlers at 22 months (Kochanska, Forman, & Coy, 1999). In addition, practices of inductive discipline, in which parents or teachers give explanations or reasons for requiring children to change behavior, are more likely to promote the children’s prosocial tendency than practices of power-assertive discipline (Hoffman, 1970, 2000).

1.1.2. Construction-based approach

According to both the psychoanalytic and learning theories, morality is the result of passive internalization. An individual’s behavior is compelled by either an internalized superego or by formed habits. In contrast, the construction-based approach does not consider morality to be an external
imposition on children. Instead, children themselves actively construct moral judgments about how people should act toward one another as children participate in social interactions with adults and peers.

In his early writings, Piaget (1932/1965) studied children’s knowledge and judgments about morality. Consistent with his general view of cognitive development, he stressed that children construct and reconstruct their moral thinking through interactions with the environment. Piaget described moral reasoning as moving from heteronomy to autonomy. In the "heteronomous" stage of moral reasoning, children strictly obey authority and adhere to prescribed rules and duties. Through overcoming egocentrism and taking perspective of others, children make progress toward the “autonomous” stage. They start to consider rules critically, and selectively apply these rules based on goals of mutual respect and cooperation.

Kohlberg (1969) modified and elaborated Piaget’s formulation, and further proposed a detailed scheme for conceptualizing and measuring moral development as a form of cognitive development. In his studies, children were asked to resolve moral dilemmas and to provide justification for their judgment. Kohlberg found a stage-like development of moral judgment. At the pre-conventional level, moral judgments are made based on obedience, punishment avoidance, and an instrumental need and exchange. At the conventional level, moral judgments are made based on role obligations and conventional conceptions of good people, such as respecting the rules and authority legitimated in the social system. At the post-conventional level, moral judgments are made beyond one’s society and justice rules, but based on contractual agreements, established procedural arrangements for resolving conflicts, mutual respect, and differentiated concepts of justice and rights.

Contemporary researchers continued this line of work by Piaget and Kohlberg, and examined children’s reasoning and the development of cognitive structures related to morality (e.g., Killen & Smetana, 2006; Rest, Narvaez, Bebeau, & Thoma, 1999). For example, Turiel and his colleagues argued that children’s judgments of moral acts are independent of conventional aspects of the social system, such as requests from authority figures or the presence of explicit rules (Nucci & Turiel, 1978; Turiel, 1983, 2006; Turiel, Killen, & Helwig, 1987). Also, Hoffman (1982) and Eisenberg (1989, 1998) proposed that
children’s perspective-taking abilities facilitate the development of prosocial tendency, such as empathy and sympathy toward others. These studies on prosociality have also been synthesized with investigations in primatology (e.g., Preston & de Waal, 2002) and neuroscience (e.g., Singer et al., 2006).

1.2. Early-emergence view

The late-acquisition view generally assumes that socio-moral expectations do not emerge until the preschool years. Recently, this assumption has been challenged by widespread speculations that children are born with a rich innate basis for understanding social interactions. This early-emergence view proposes that moral development builds on early-emerging socio-moral intuitions about how individuals should act toward one another (e.g., Boyd & Richerson, 2005; Dupoux & Jacob, 2007; Dwyer, 2007; Greene, 2005; Haidt, 2008; Jackendoff, 2007; Mikhail, 2007; Premack, 2007; Sigmund, Fehr, & Novak, 2002; Wynn, 2008).

Researchers with the early-emergence view argue that the human brain is programmed to make it easy to learn certain things and difficult to learn other things. As Marcus (2004) described, “The initial organization of the brain does not rely that much on experience… Nature provides a first draft, which experience then revises…‘Built-in’ does not mean unmalleable; it means organized in advance of experience.” (Marcus, 2004, p. 12)

This early-emergence view can be further divided into two major approaches. One approach emphasizes the role of emotions in the generation and development of moral judgments (emotion-based approach). The other approach emphasizes the role of tacit principles in children’s reasoning about social interactions (principle-based approach).

1.2.1. Emotion-based approach

As discussed above, the construction-based approach (Kohlberg, 1969; Piaget, 1932; Turiel, 1983) argues that moral knowledge and moral judgments are reached primarily by a process of conscious, language-based reasoning. However, some researchers (e.g., Haidt, 2001) propose that moral emotions, or “gut feelings,” actually come first. These feelings directly cause post-hoc constructions, which people generate to justify their implicit and automatic emotions. Haidt’s argument was built on philosophical
claims by Hume who considered moral reasoning to be “the slave of the passions.” Hume said that “the ultimate ends of human actions can never be accounted for by reason, but recommend themselves entirely to the sentiments and affections of mankind” (1777/1966, p. 131).

Therefore, if moral emotions drive moral reasoning, then moral judgment is primarily a bottom-up process that is triggered by external stimuli rather than by concepts or principles in one’s mind. For example, in a study (Haidt, Bjorklund, & Murphy, 2000), upon hearing a story that a brother and a sister made love during a summer vacation, most participants immediately felt disgusted and said that it was wrong. If asked why they felt this way, they then began searching for reasons. They pointed out the dangers of inbreeding, even though birth control was implemented; they argued that the siblings would be hurt, perhaps emotionally, even though the story made it clear that no harm befell them.

1.2.2. Principle-based approach

In contrast to the emotion-based approach, which advocates the causal role of emotions in generating moral intuitions, the principle-based approach argues that emotions are the byproduct of moral judgments. According to the latter, infants are born with tacit principles dedicated specially for socio-moral reasoning. These principles are deployed without awareness and can hardly be explicitly articulated. Two accounts of the principle-based approach have been offered: the “universal moral grammar” account and the “moral causal framework” account.

The “universal moral grammar” account (e.g., Mikhail, 2007; Raw, 1971) was developed based on an analogy between the study of moral cognition and Chomsky’s generative linguistics research program (e.g., Chomsky, 1965; 1975). Chomsky argues that people’s language comprehension and production abilities rely on their tacit knowledge of a grammatical system, of which they are not aware. The grammatical system is a network of computational principles connected to a finite set of binary parameters (similar to binary switches). In each human language, the parameters are instantiated with a particular setting. With this built-in computational system, children can easily learn their local languages from a finite and fragmentary sample of linguistic input.
Similar to grammatical judgments, moral judgments may result from complex and unconscious computations involving abstract computational principles. For example, in some studies using the “trolley dilemmas” (e.g., Thomson, 1971), participants were presented with scenarios in which a bystander sacrificed one person to save five, and all scenarios ended with the same number of casualties. The participants considered it more acceptable for a bystander to sacrifice that person when the person’s death was merely a foreseen consequence than when it was an intended consequence of the bystander’s act. This minimal contrast between trolley dilemmas seemed to have been taken into account when the participants unconsciously computed their moral judgments.

However, the analogy between morality and linguistics may not be feasible. Speakers of two distinct languages can hold irreconcilably divergent grammatical intuitions (e.g., whether or not nouns precede adjectives). But many moral reactions are widely shared among members of various cultures (Rozin et al., 1999; Shweder et al., 1995). For example, most cultures consider murder and rape as unacceptable. The similarities of moral intuitions across cultures cannot be derived from the arbitrary outcomes of the culture-dependent parameters in the moral grammar system, but they are more likely to reflect a small set of universal moral norms, with which humans are equipped at birth.

To better interpret the cross-cultural similarity of moral intuition, the “moral causal framework” account proposed that infants are born with a skeletal causal framework for interpreting social interactions among individuals (e.g., Dupoux & Jacob, 2007; Dwyer, 2007; Geraci & Surian, in press; Hamlin, Wynn, Bloom, 2007; Jackendoff, 2007; Olson & Spelke, 2008; Premack, 2007; Wynn, 2008). This system operates based on a small set of universal moral principles. Moral diversity across cultures may due to different moral adjudication among competing moral intuitions. Each culture has its own elaboration and ranking of members of a finite set of universal moral principles (Dupoux & Jacob, 2007). For example, some individualistic cultures put more weight on individual’s privacy than on ingroup loyalty, while other cultures may show the reverse preference ordering of the two norms.
2. Documented socio-moral norms

Researchers have started to look at which socio-moral norms might emerge early in life. Below are some examples of moral norms that researchers have investigated.

2.1. Norm of reciprocity

According to the norm of reciprocity, if A acts in some way toward B, and B chooses to respond, then B’s action should match that of A in value, though not necessarily match in form. The norm of reciprocity can be separated into two corollaries: reciprocation, stating that B should act positively (or reciprocate) in response to a positive action, and retaliation, stating that B should act negatively (or retaliate) in response to a negative action. Reciprocity specifies how an individual should act in response to another’s action, and is as such conceptually distinct from altruism and other prosocial behaviors (e.g., Eisenberg, Fabes, & Spinrad, 2006).

Researchers from diverse fields—including anthropology, economics, linguistics, psychology, and sociology—have argued that reciprocity is one of the fundamental norms guiding human social interactions (e.g., Fehr & Gachter, 2000; Fry, 2006; Gouldner, 1960; Jackendoff, 2007; Premack, 1990). Adults possess an abstract calculating system that describes the effect of an individual’s action upon another individual (e.g. Jackendoff, 2006). This value system is composed of two dimensions: valence and magnitude. Valence describes whether an individual’s action exerts a positive, negative, or neutral effect on others, while magnitude describes how strong the effect is. By combining the concepts of valence and magnitude, people reason about individuals’ actions and interactions. Jackendoff (2006) studied the acceptability of linguistic expressions concerning valence and magnitude in social interaction. For example, people expect a negative action in return for a negative action, as in the scenario “Fred slashed Lois’s tires for insulting his sister”. On the other hand, people find it absurd if Lois’s initial action and Fred’s reaction do not match in valence. For example, “Fred cooked Lois dinner for insulting his sister”. In addition, people find it odd if Lois’s initial action and Fred’s reaction mismatch in magnitude even though they match in valence. The scenario “Fred slashed Lois’s tires for eating too little at dinner” is an over-reaction, while the scenario “Fred slashed Lois’s tires for murdering his entire family” is an
under-reaction (Jackendoff, 2006, p. 395).

Beginning with Piaget’s seminal work (1932/1965), developmental psychologists have long been interested in determining at what age and under what conditions children demonstrate an expectation of reciprocity. Investigations using first-party situations examine whether children show reciprocity in their own responses to others’ actions. Investigations using third-party situations test whether children appeal to reciprocity when interpreting or evaluating others’ actions. There is considerable evidence, from various first- and third-party investigations, that children aged 6 years and older understand and adhere to the norm of reciprocity (e.g., Berndt, 1977; Dahlman, Ljungqvist, & Johannesson, 2007; Darley, Klosson, & Zanna, 1978; Staub & Sherk, 1970; Suls, Witenberg, & Gutkin, 1981).

Studies with 3- to 5-year-olds have also yielded a number of positive findings. First, naturalistic observations in preschool settings indicate that young children show reciprocity when interacting with peers (e.g., Fujisawa, Kutsukake, & Hasegawa, 2008; Leiter, 1977; Strayer, Wareing, & Rushton, 1979). For example, Fujisawa et al. (2008) observed Japanese 3- and 4-year-olds at their nursery school for a year and examined how often they reciprocated two positive actions, helping and object offering, during free-play time. At each age, children reciprocated both behaviors; helping was reciprocated about equally with friends and non-friends, but object offering was reciprocated more with friends. Second, first-party experimental tasks with young children have also produced encouraging results (e.g., Dahlman et al., 2007; Levitt, Weber, Clark, & McDonnell, 1985). For example, Dahlman et al. (2007) divided 3-to 5-year-old Swedish children into anonymous pairs of givers and recipients. During the first phase of the procedure, the giver chose between giving the recipient a reward (a bag of raisins) or no reward. During the second phase, the recipient chose between the same two options for the giver. Recipients were somewhat more likely to give a reward if they had received one in the first phase than if they had received none. Finally, some evidence of sensitivity to reciprocity was also found in a third-party task with 3- to 5-year-olds (Berndt, 1979): children were more likely to invoke retaliation to explain cartoon characters’ negative actions when provoked. Therefore, children by 3 to 5 years of age show sensitivity to the norm of reciprocity in some first- and third-party situations. The question is whether this sensitivity is already
present in infancy, or if it emerges in the toddler or preschool years.

Some evidence shows that infants assign valences (positive, neutral, or negative) to actions when watching two individuals interact (e.g. Hamlin et al., 2007; Kuhlmeier et al., 2003; Premack, 1990; Premack & Premack, 1997). For example, in the experiments by Premack and Premack (1997), one-year-old infants were first habituated to one of four events involving two self-propelled agents, a grey and a black ball: the hit-vertical (i.e. the grey ball pursued and struck the black ball vertically several times), the caress (i.e. the grey ball had gentle contacts with the black ball), the help (i.e. the grey ball helped the black ball get over a barrier), and the hinder event (i.e. the grey ball prevented the black ball from getting over a barrier). Next, all the infants were presented with the same test event: a hit-horizontal event (i.e. the grey ball pursued and struck the black ball horizontally several times). The infants who were habituated to a positive event (i.e., help or caress) showed a substantial dishabituation to the negative test event, but those who were habituated to a negative event (i.e., hinder or hit-vertical) showed little or no dishabituation to the negative test event. These findings suggested that infants discriminate between actions directed at others that have a positive (beneficial) valence and those that have a negative (detrimental) valence. Specially, they view caressing and helping as positive but hitting and hindering as negative.

Infants not only distinguish valences of actions, but also understand that valences have consequences for subsequent interactions. Infants prefer—and expect others to prefer—individuals who produce positive as opposed to negative actions (e.g., Hamlin, Wynn, & Bloom, 2007). Moreover, infants tend to act positively in response to another individual’s well-intended actions (Dunfield & Kuhlmeier, 2010). For example, infants at the age of 21 months selectively help those who have shown a willingness to provide. They preferred to help an individual who previously intended to provide a toy (either with or without successful outcomes) over an individual who was unwilling to provide. Similarly, toddlers selectively avoid helping people with harmful intention (Vaish, Carpenter, & Tomasello, 2010). 3-year-olds watched an adult harm another adult. Subsequently, the children helped the harmful adult less often than a third (previously neutral) adult; also, they helped an actor who intended but failed to harm another
adult less often than a neutral adult.

Prior study also suggests that infants expect individuals to reciprocate or retaliate in accordance with the norm of reciprocity (He, Bolz, Baillargeon, Premack, under review). In the experiments, 15-month-old infants watched a live two-part vignette involving two unfamiliar female experimenters, A and B. In the first part of each vignette, A acted either positively or negatively toward B; in the second part, B now acted either positively or negatively toward A. Infants detected a violation (as indicated by longer looking times) when B acted negatively in response to A’s positive action, or when B acted positively in response to A’s negative action. For example, after seeing A steal B’s cookie, infants were surprised if B helped A put away her sticker, but not if B tore up A’s sticker. These results indicate that infants in their second year of life already hold a rudimentary tit-for-tat stance on social interactions: infants expect reciprocal actions to match initial actions in valence, but recognize that they need not match in form.

2.2. Norm of ingroup loyalty

The norm of ingroup loyalty states that individuals in a social group tend to act in ways that support the group. For example, individuals should prefer ingroup over outgroup members, should provide assistance and cooperation to ingroup members, and should act to maintain positivity and harmony within groups (e.g., Brewer, 1999; Haidt & Joseph, 2007; Kinzler et al., 2007, 2011; Mahajan & Wynn, 2010; Shutts et al., 2009; Werneken & Tomasello, 2006, 2007).

Some evidence suggests that infants are already sensitive to social groups. Infants appear to be sensitive to ways of aggregating individuals by race, gender, age, language, shared preference and so on. They seem to notice various dimensions of similarities and differences among people, which signal different social groups. Based on these dimensions, infants form a preference for ingroup over outgroup members.

For example, infants as young as 3 months prefer to look at faces of their own race over those of a different race (e.g. Kelly et al., 2005). This visual preference based on racial differences is not present in the first days of life, but is learned and derived from differences in exposure to own- versus other-race faces within the first three months of life. These findings are consistent with those from studies examining
the influence of infants’ prototypical racial environment on infants’ racial preference. Evidence shows that infants prefer the race that they have more exposure to. For example, Israeli infants from Caucasian families prefer to look at Caucasian over African faces, and Ethiopian infants from African families prefer to look at African over Caucasian faces. However, Israeli infants from African families, living in a predominantly Caucasian culture, show no consistent preferences (Bar-Haim et al., 2006). Moreover, infants become more accurate at processing faces of their own race than those of other races before their first birthday (e.g. Kelly, et al, 2007; Anzures, et al, 2009), a bias that is maintained during later childhood (e.g. Shutts & Kinzler, 2007).

Findings from studies on infants’ gender preference are also consistent with those on racial preference. For example, infants aged 3 to 4 months prefer to look at faces of the same gender as their primary caregiver. That is, regardless of whether they themselves are male or female, infants reared with female primary caregivers tend to show a preference for female faces, but infants reared with male primary caregivers prefer male faces (e.g. Quinn et al., 2002). In addition, 12-month old children tend to choose dolls that are most like the children themselves in gender and in race (e.g. Katz, 2002).

Besides gender and race, age is another powerful guide for young children’s social preferences. For example, 3-year-old children show a bias toward people of their own age. If 3-year-old boys were asked to choose between two toys, one offered by another boy and one offered by a male adult, they were more likely to choose the toy from the boy (e.g. Shutts, Banaji, & Spelke, 2009).

Infants also prefer people who speak the language of their own speech community. For example, 6-month-old infants from mono-lingual American English family watched alternating films showing two adult women speaking in American English and Spanish (both women were bilingual so that each of them spoke Spanish to half the infants and English to the others). The women were then presented side by side, smiling but silent. The infants looked longer at the woman who had spoken to them in English, their native language (Kinzler, Dupoux, & Spelke, 2007). A similar preference was found even in neonates (Mehler et al., 1993). What’s more, variations of accents can also make a difference: 5-month-olds prefer looking at individuals who previously spoke their native language with a native accent as opposed to a
foreign accent. With age, they show a more explicit social preference beyond their visual preference: 10-month-olds reach for toys that are offered by a speaker of their native language rather than by a speaker of a foreign language (Kinzler, Dupoux, & Spelke, 2007).

In addition, infants tend to act more positive toward people with whom they share preferences. In a series of studies, Mahajan and Wynn (2010) demonstrated that, like adults, infants 7 to 14 months of age prefer to interact with those who share trivial similarities to themselves. Specifically, infants who choose between two items (two distinct snacks, toys, or mittens) prefer individuals who make the same choice over those who choose the opposite. Infants’ performance in a control condition also suggests that these preferences are based on a cognitive comparison process (‘‘like me’’/’’different from me’’), rather than reflecting associative preferences.

Finally, prosocial behaviors such as helping, which are tied to ingroup loyalty, can be manipulated by merely priming with affiliation. In a study by Over and Carpenter (2009), 18-month-old infants’ helping behavior was measured after being primed with photographs in one of the four conditions: in the affiliation condition, two small dolls stood next to each other in the background of a familiar household object; in the individual condition, a single doll stood alone in the background of otherwise identical photographs; in the neutral, non-social condition, two small stacks of blocks were placed in the background; and in the non-affiliative relation condition, the two dolls stood back to back. All the photographs had similar foregrounds, but their backgrounds varied depending on the condition. The infants provided assistance to an unfamiliar experimenter more often and more spontaneously in the affiliation condition than in the other three conditions. Hence, infants’ prosocial behavior can be promoted if their mindset of ingroup loyalty is activated.

2.3. Other norms

In addition to the norms of reciprocity and ingroup loyalty, researchers have also investigated a number of other socio-moral norms, such as the norms of fairness, welfare, and hierarchy.

The norm of fairness states that individuals should deal fairly with others. Research on fairness in adults from Western cultures have shown that adults are more interested in fair play than in self-gain (e.g.,
Camerer, 2003; Guth, Schmittberger, & Scharze, 1982; Henrich et al., 2001; Sigmund, Fehr, & Novak, 2002). Similar results have been observed in different cultures around the world, leading researchers to speculate that a concern for fairness may be innate in humans, and may have evolved during the millions of years that our ancestors lived in small bands of hunter-gatherers, when survival depended on the fair sharing of resources and efforts (e.g., Boyd & Richerson, 2005; Fehr & Fischbacher, 2003; Novak, Page, & Sigmund, 2000; Premack & Premack, 2003). Similarly, children also demonstrate a concern for fairness (e.g., Fehr, Bernhard, & Rockenbach, 2008; Olson & Spelke, 2008; Sutter, 2007; Turiel, 1983).

For example, in a study by Olson and Speke (2008), 3.5-year-old children were presented with five dolls: one was identified as the protagonist and the other four were identified either as the protagonist’s siblings, as the protagonist’s friends, or as strangers. The children were asked to help the protagonist distribute resources to the other dolls (e.g., divide two, three or four stickers among two siblings and two strangers). When the number of resources and recipients were equal, children directed the protagonist to share resources equally with all potential recipients; but when the number of available resources was smaller than the number of recipients, children directed the protagonist to share more with family and friends than with strangers (i.e., consistent with the norm of ingroup loyalty).

Recent research found that even infants have some rudimentary expectations that individuals should act fairly when distributing resources between other individuals or when rewarding others for their efforts. For example, in a recent study by Geraci and Surian (in press), 16-month-olds first saw a “fair distributor” perform equal distributions toward two recipients and an “unfair distributor” perform unequal distributions toward another recipients, while a bystander observed all the distributive actions. Infants looked reliably longer when the bystander approached the “fair distributor” than when the bystander approached the “unfair distributor.” Infants’ manual choices between the two distributors also reveal their preference toward the “fair distributor” over the “unfair distributor.” Moreover, in the experiments of Sloane, Baillargeon, and Premack (under review), 19-month-old infants viewed events in which an experimenter distributed two toys unfairly between two giraffe puppets as unexpected.
The norm of welfare states that individuals should refrain from doing serious harm to others and should help others who are in distress (e.g., Jackendoff, 2009; Kuhlmeier et al., 2003; Premack & Premack, 1995, 1997). The norm of welfare is violated by inflicting harm on another individual (i.e., an error of commission), or by ignoring another individual who is in distress and needs help (i.e., an error of omission). Some evidence indicates that infants, by their second year of life, may already possess expectations about how individuals should respond to others who are in distress (e.g., Johnson, Dweck, & Chen, 2007; Johnson, et al., 2010; Houston, et al., 2010). For example, when observing a computer-animated scene involving a “caregiver” and a crying “child” (depicted by a large and a small oval, respectively), infants detect a violation if the caregiver ignores, rather than comforts, the distressed child (Johnson, et al., 2007).

The norm of hierarchy states that dominant individuals have privileged influence and access to resources compared to subordinate ones. Dominance hierarchies are ubiquitous across cultures (e.g., Fiske, 2000). A recent study (Thomsen et al., 2011) found that infants, by one year old, realize social dominance and understand that when two individuals' goals conflict, the one with larger relative size will prevail. Infants watched videos of interactions between a large and a small cartoon block, each depicted with eyes and a mouth. In the familiarization trials, either the large block bounced across a stage in one direction or the small block bounced across a stage in the opposite direction. Next, infants watched the two blocks meet in the middle, impeding one another's progress. They saw either the large block deferred the small one by bowing and stepping aside, or the small block deferred the large block by bowing and stepping aside. Infants aged 10 to 16 months looked significantly longer when the large block yielded to the smaller one than when the small one yielded to the large one. This and control results together suggest that infants already have some rudimentary understanding of social dominance and how it relates to relative size.
3. Present research

3.1. Overview of present projects

The present projects built on the previous findings and continued to examine socio-moral expectations in infants and toddlers. In particular, we asked whether children’s expectations about reciprocity would be modulated by considerations of ingroup loyalty, and if so, then how.

The first two projects were aimed at testing whether infants’ expectations of retaliation would be modulated by considerations of ingroup loyalty.

In Project 1, we tested the idea of “escalation” and asked whether infants would expect an individual to retaliate more severely against an outgroup member than against an ingroup member.

In Project 2, we tested the idea of “co-retaliation” and asked whether infants would expect individuals to retaliate against an outgroup member who had previously hindered the individual’s group member. We also examined the continuity of early socio-moral expectations by testing “co-retaliation” in toddlers using similar methods.

In Project 3, we started to investigate whether young children’s expectations of reciprocation would also be modulated by considerations of ingroup loyalty. In particular, we tested the idea of “privilege” and asked whether toddlers would expect individuals to act more positively in response to a friendly overture by an ingroup member than by an outgroup member.

To sum up, we explored whether infants and toddlers would have early expectations of socio-moral norms. We reasoned that the findings of the present projects would help shed light on the origin of socio-moral expectations. Given that infants have very limited social experience besides the child-caregiver interactions, we reasoned that positive findings in the present projects would support the early-emergence view that children are born with a rich innate basis for understanding social interactions.

3.2. Features of the research approach

Here are three key features of our experimental approach: we presented infants and toddlers with third-party situations involving novel artificial social groups and we examined their sensitivity to contextual factors of the situations.
3.2.1. Using novel artificial social groups

The present projects used arbitrarily-defined social groups children had no prior exposure to. Most prior research has used socially-recognized groups, such as race, gender, nationality, religion, and language groups, which children may have exposure to in daily life. It would be difficult to disentangle whether children’s expectations about interactions among these socially-recognized groups is due to familiarity through experience or whether it is due to general experiences of how people should act within or across groups. By using novel artificial social groups, we could limit the effect of children’s experience of socially-recognized groups.

This feature allowed us to examine whether a mere categorical distinction, prior to any cultural elaboration, would have some consequences for children’s expectations about how individuals should interact with one another. The minimal-group paradigm was developed and used to study the origins of intergroup bias in adults (Billig & Tajfel, 1973; Brewer, 1979; Karp, Jin, Yamagishi, & Shinotsuka, 1993; Tajfel, Billig, Bundy, & Flament, 1971). In the studies by Tajfel and his colleagues (e.g., Billig & Tajfel, 1973; Tajfel et al., 1971), adult participants were grouped based on shared preferences (e.g., whether they prefer Klee’s or Kandinsky’s paintings) or even random assignment (e.g., participants were told whether they under- or over-estimated the number of dots in a dot estimation task). These seemingly meaningless social groupings were found to be sufficient to induce preference for the ingroup across a wide range of measures, including resource allocation (Locksley, Ortiz, & Hepburn, 1980; Tajfel & Turner, 2004), trait evaluations (Locksley, Ortiz, & Hepburn, 1980), as well as implicit measures designed to tap introspectively inaccessible associations in semantic memory (Ashburn-Nardo, Voils, & Monteith, 2001; Otten & Wentura, 1999).

In a series of studies by Bigler and colleagues (Bigler, Brown, & Markell, 2001; Bigler, Jones, & Loblinier, 1997; Patterson & Bigler, 2006), children were assigned to novel groups (e.g., blue- or yellow-shirt team) over the course of a 3 to 6 week summer program. The groups were used by teachers for different functions, for example, as a way of dividing the children up for activities. The results showed
that the intergroup categorization had significant effect on both preschool and elementary-school children’s social judgment regarding those novel groups.

The present projects examined whether infants and toddlers would consider unfamiliar, arbitrarily-defined group markers, such as novel physical appearance or naming label, to be sufficient to establish group membership. If infants and toddlers are sensitive to novel markers for social groups and immediately have expectations about how individuals should act within or across groups, it would support the early-emergence view that children’s moral sense is either innate or learned quickly in infancy.

3.2.2. Using third-party situations

Most previous research on social groups has used first-party situations, where children choose to interact with either a member of their own group or one of another group. However, first-party situations may impair children’s moral reasoning because children’s self-interest is closely related to the outcomes of the situations. Here we used third-party situations. Instead of examining how infants and toddlers themselves act towards their ingroup and outgroup members, we asked a more basic question: when children observe individuals’ interactions as bystanders, what kinds of expectations do they have for these interactions within and across groups?

In the present projects, children themselves did not benefit or incur harm from the possible outcomes of the situations; and they were also not affiliated with any of the social groups. Therefore, we minimized the involvement of children’s self-interests in the events that they were watching. Young children are generally poor at regulating their emotions or controlling their behaviors when their self-interest is closely involved (e.g., Damon, 1975; Fehr et al., 2008; Piaget, 1932/1965). By placing children in the third-party role, they would be more likely to reason about how people should act in a general sense, rather than how they themselves would act in that situation.

By using third-party situations, we measured infants and toddlers’ spontaneous non-verbal responses to the situations. Evidence that infants and toddlers reveal their moral sense in non-verbal measures would be important for one crucial methodological reason; it would broaden the assortment of experimental tools available for exploring early moral development. Prior research on moral judgments or
moral reasoning has relied heavily on participants’ *verbal* responses to complicated stories that often involve some sort of dilemma (such as the trolley dilemmas). Infants and toddlers are either unable to, or have minimal practice at articulating what they think is right or wrong. However, they are well-versed in exploring the world with their hands and eyes. Research using non-verbal tasks (e.g., measuring what children reach for, where they look, and how long they look) has already shown that infants come into the world equipped with a wealth of cognitive resources to support their reasoning about their surroundings. For example, infants prefer human faces (e.g. Fantz, 1958); infants can represent hidden objects (e.g. Baillargeon, Spelke, and Wasserman, 1985); infants have some expectations about when an object will fall and when it will be stable (e.g. Baillargeon, Needham, & DeVos, 1992); infants can segregate objects based on objects’ featural differences (e.g. Xu, & Carey, 1996; Needham & Baillargeon, 2000); infants can infer the location of an object based on its height, width or other variables (e.g. Hespos & Baillargeon, 2006); infants can understand that others may have and act on their beliefs, either true or false (e.g., Baillargeon, Scott, & He, 2010; He, Bolz, & Baillargeon, 2011; Onishi & Baillargeon, 2005). Using similar non-verbal measures, the present projects aimed at exploring potential early socio-moral expectations.

### 3.2.3. Exploring sensitivity to contextual factors

Prior research has found that adults’ emotions and their consideration of contextual factors make *independent* contributions to moral judgments. For instance, victims of natural disasters often lead us to feel personal distress and concern, but do *not* lead us to judge that a moral transgression has occurred. Similarly, individuals respond affectively but *without* drawing any moral judgment when other people’s suffering is a result of an accident. Moral judgments are more often made when a person is suffering because of an unjust violation of socio-moral norms (Nichols, 2008).

Here we examined young children’s sensitivity to contextual factors when generating socio-moral expectations, in other words, whether they can take into account not only the outcomes of actions, but also the contexts in which the actions occur. The present projects focused on three main types of contextual factors: (1) whether young children would consider the *intergroup context* (e.g., the individuals
belong to the same group or different groups); (2) whether young children would consider the prior history of interactions among the individuals (e.g., the presence of prior provocation, the previous actions targeted toward the individual’s ingroup members as well as those targeted toward the individual him/herself); (3) whether young children would attribute differentiated mental states to each member of a group (e.g., the individuals’ motivations and intentions; their knowledge and ignorance; and their false beliefs and pretense).

Evidence for early sensitivity to these contextual factors would support the early-emergence view, and would be especially consistent with the principle-based approach, as opposed to the emotions-based approach. The emotion-based approach predicts that early moral expectations should be primarily triggered by external emotional stimuli, such as how violent actions are (a bottom-up process); in contrast, the principle-based approach argues that early moral expectations are influenced by internal moral principles that are applied based on particular contextual factors (a top-down process).
CHAPTER 2: PROJECT 1 ON ESCALATION

Project 1 examines the origins of the norm of reciprocity, specifically the aspect of retaliation. Several key points regarding this phenomenon have already been established. First, retaliatory behavior is evident in both human (e.g., Bernhard, Fischbacher, & Fehr, 2006; Fehr & Gachter, 2002; Shinada, Yamagishi, & Ohmura, 2004) and non-human societies (Boyd et al., 2003; Clutton-Brock & Parker, 1995; Clutton-Brock, Price, & MacColl, 1992; de Waal, 1982; Jensen, Call, & Tomasello, 2007; Manson, 1994; Nadler & Miller, 1982; Smuts & Smuts, 1983).

Second, retaliation is not rare within a social group. When a family member inflicts physical violence, when a partner commits adultery, or when a tribe member cheats in a social exchange, their bad behavior often evokes paybacks. People have a sense of satisfaction after extracting vengeance or getting even (e.g., Fry, 2006).

Third, retaliation is also used as a way to regulate the interactions across groups. For example, among the tribes of the Upper Xingu River basin in Amazonia, there exists a regular intertribal ritual called a Yawari contest: two members from opposed tribes pair off, then hurl insults and subsequently wax-tipped spears at one another. All the spears are required to be tipped with wax and they are not allowed to be thrown above the waist. People have to control their aggressive impulse because the aggressors and their targets change roles in the ritual, which means if a person inflicts too much harm, he may find himself equally injured shortly (Gregor, 1994).

While there is emerging evidence that even infants are aware of and appear to expect retaliation in certain situations (e.g., He, et al, under review), many factors are still unexplored. It is unclear whether infants expect individuals to follow the norm of reciprocity in both the ingroup and the outgroup context, and if so, to what extent (magnitude).

Some evidence suggested that the ingroup feature of friendship affects the quantitative aspect of reciprocity. For example, one observational study (Fujisawa, Kutsukake, & Hasegawa, 2010) recorded and analyzed reciprocal behaviors among 3- and 4-year-old Japanese preschool children during free-play time. The results show a tendency for positive reciprocity. Both friends and non-friends tended to
reciprocate the actions of those who had previously offered help compared to those who had not, but friends reciprocated object offerings more frequently than non-friends. However, it is unclear whether friendship would affect expectations about retaliation (i.e., negative reciprocity) in a similar manner.

The properties of negative reciprocity have been explored in adults. Some studies have examined whether relatedness to a perpetrator affects people’s willingness to punish the perpetrator (e.g., Boyd et al., 2003; O’ Gorman, Wilson, & Miller, 2005). The study by O’ Gorman et al. (2005) showed that participants assigned equivalent levels of punishment to a cousin, friend, and stranger who stole from a group account. The lack of significant differences among the perpetrators was not due to failure of manipulating the variable of relatedness; the same manipulation yielded significant results with helping behavior -- participants were more willing to provide assistance to cousins than strangers. While it seems that the participants would like to retaliate against ingroup members to a similar extent as against strangers, this may be specific to the “public goods” context used in this experiment. This has been shown to follow different psychological principles compared to non-collective interactions (e.g., Lieberman & Linke, 2007). In addition, how infants would react to such a situation has also not yet been investigated. Given the early sensitivity to ingroup loyalty (e.g., Over & Carpenter, 2009), infants may expect a more severe retaliation against an outgroup member than against an ingroup member.

Another intriguing question is whether children would expect any limit on the severity of retaliation against an outgroup member. This might be plausible for a number of reasons.

(1) The severity of punishment for a misdeed or wrongdoing should be reasonable and proportional to the severity of the infraction, regardless of whether the target of the punishment is an ingroup or an outgroup member (e.g., Fry, 2006). The expressions of “an eye for an eye” and “a tooth for a tooth” reflect the balance of negative paybacks. A serious misdeed tends to elicit a greater payback than a minor breach of rules or social obligations (“let the punishment fit the crime”).

(2) The norm of “do no harm” (e.g. Premack, 2007; Haidt & Joseph, 2007) may also limit an extremely severe retaliation against an outgroup member, even when escalation is expected. Any extreme retaliatory action, such as taking a life in exchange for a life, is generally disapproved. For example, one
study sampled 50 cultures and found the occurrence of lethal feuding to be as follows: frequent in 16%; infrequent in 28%; and absent in the remaining 56% (Otterbein & Otterbein, 1965).

(3) The boundaries among groups are dynamic and overlapping. People frequently trade goods and seek marriage across group lines. Today’s outgroup member could be your ingroup member tomorrow. Therefore, extreme retaliation may be abandoned because of the potential for merged or swapped group membership in the future.

Based on the findings of these previous studies, we generated 3 hypotheses to investigate in Project 1: (1) Infants expect individuals to retaliate against ingroup as well as against outgroup members; (2) Infants expect individuals to retaliate more severely against outgroup than against ingroup members. In other words, infants expect an escalation of retaliation against outgroup compared to ingroup members; and (3) infants expect a limit on the severity of retaliation even against outgroup member. In summary, infants expect the norm of retaliation in both ingroup and outgroup contexts, and considerations of ingroup loyalty will modulate the magnitude, but not the valence of retaliatory behaviors.

One point to take into account while exploring these ideas is that infants might expect more retaliation against outgroup members in part because of different set-points: if some degree of aggression toward outgroup members is tolerated without provocation, it would help explain why more aggressive reactions are expected with provocation. To address this issue, we included a no-provocation condition to examine whether children would view it as acceptable to inflict aggression on an ingroup or an outgroup member without provocation.

In Project 1, infants were presented with two arbitrary social groups marked by contrasting outfits. This allowed us to examine whether young children would form social groups based on surface features of the individuals, such as their physical similarity. People tend to expect physically alike individuals to belong to the same groups and physically unlike ones belong to different groups. Prior research showed that children can categorize people into arbitrary social groups based on physical similarity. For example, in the experiments of Olson and her colleagues (Olson, Banaji, Dweck, & Spelke, 2006), children aged 3-5 years showed a preference for lucky over unlucky people, and generalized this preference to the lucky
people’s ingroup members who wore T-shirts of the same color, despite the fact that the color groups were never explicitly mentioned. Therefore, we speculated that infants might be able to categorize people into groups based on salient physical similarities and differences.

In Experiment 1, we tested whether infants would expect an individual to retaliate against an ingroup member, and compared between the magnitude of the negative action expected with and without provocation.

In Experiment 2, we tested whether infants would expect an individual to retaliate against an outgroup member, and we again compared between the magnitude of the negative action expected with and without provocation. Of particular interest was whether infants would expect a more severe retaliation against an outgroup than an ingroup member.

In Experiment 3, we tested whether infants would expect any limit on the severity of the retaliation allowed against an outgroup member.

**Experiment 1**

Prior experiments suggest that infants have an expectation of retaliation (e.g., He, et al, under review), but the group membership of the two individuals was not manipulated in these experiments. Experiment 1 tested (1) whether infants expect an individual to retaliate against an ingroup member, (2) whether the magnitude of the negative action expected following provocation is greater than that expected when no provocation has occurred.

Infants watched one familiarization and one test trial (see figure 1). At the start of the familiarization trial, three individuals sat at an apparatus. A1, on the right, and B1, on the left, wore different outfits and thus belonged to different groups; the target individual (who retaliated against A1 in the test trial) sat in the middle and wore the same outfit as A1. In the familiarization trial, as B1 and the target individual looked on, A1 stacked small blocks and built two towers of four blocks each.

In the test trial in the provocation situation, B1 was absent, and the target individual ate crackers as she watched A1 build a third tower of blocks. After A1 completed her tower, she stole one of the target individual’s crackers and left. In retaliation, the target individual knocked down either one block from one
tower (one-block event), or one or two towers (some-towers event); the target individual then paused until the trial ended. The test trial in the no-provocation situation was similar except that, after A1 completed her tower, she left without stealing the target individual’s cracker.

We predicted that infants would expect the target individual to act more negatively toward her ingroup member when provoked than when unprovoked.

Method

Participants

Our design called for 32 participants. Thus far, 28 healthy term infants (13 boys and 15 girls) have participated in this experiment (17 months, 9 day to 19 months, 23 days, $M = 18$ months, 21 days). Another 10 infants were tested but excluded, three because they were fussy (1), distracted (1), or active (1), and seven because they looked for maximum allowed in test. Sixteen infants were randomly assigned to the provocation ($M = 19$ months, 2 days) and the no-provocation ($M = 18$ months, 11 days) situations. In each situation, half the infants saw the one-block event, and half saw the some-towers event.

Participants in this and the following experiments were recruited primarily from purchased mailing lists and from birth announcements in the local newspaper. Parents were offered reimbursement for their transportation expenses but were not compensated for their participation.

Apparatus and Stimuli

The apparatus consisted of a brightly lit wooden display booth (124 cm high X 102 cm wide X 56.5 cm deep) mounted 77 cm above the floor in a brightly lit test room. The infant sat on a parent’s lap and faced a large opening (46 cm X 95 cm wide) in the front of the apparatus; between trials, a curtain consisting of a muslin-covered wooden frame (59.5 cm X 101.5 cm) was lowered to hide this opening. Two large cloth-covered wooden frames (each 183 cm X 76 cm) stood at an angle on either side of the apparatus in order to isolate the infant from the test room. Inside the apparatus, the side walls were painted white, the back wall was made of white foam board, and the floor was covered with granite-patterned contact paper.
Three female experimenters were involved in the events. A1 knelt at a window (51 cm X 38 cm) in the right wall of the apparatus, according to the infant’s perspective. This right window was located 4.5 cm above the apparatus floor and 7.5 cm from the back wall. A1 wore a sleeveless orange giraffe-patterned smock with a white long-sleeve shirt underneath. Her hair was tied back and she wore a blue hat decorated with white fringes along the center. B1 knelt at a window (50 cm X 37.5 cm) in the left wall of the apparatus. This left window was located 4.5 cm above the apparatus floor and 7.5 cm from the back wall. B1 wore a long-sleeve shirt made of brown fabric resembling bear fur with a black long-sleeve turtleneck underneath. Her hair was tied back and she wore a red knit cap with alternating green and yellow pompoms along the bottom. The target individual knelt at a window (50.5 cm X 54.5 cm) in the back wall of the apparatus. This back window extended from the apparatus floor and was located 24 cm from the right wall. The target individual wore the same kind of outfit as A1. White curtains behind the three experimenters hid the test room.

During the familiarization trial, A1 brought in a green plastic woven basket (8 cm X 14 cm X 11 cm at its largest points) containing 12 alphabet blocks (each 3 cm X 3 cm X 3 cm). Four blocks each were painted with a large blue letter D, four with a large yellow letter C, and four with a large red letter F.

During the test trial, a plate (a black shallow plastic dish, 0.5 cm X 10 cm in diameter, located 2 cm from the back wall and 44 cm from the right wall) of 20 Pepperidge Farm® Goldfish Crackers (0.5 cm X 1.5 cm X 1.5 cm at its largest points) was positioned in front of the target individual.

Trials

In the following descriptions, the numbers in parentheses indicate the number of seconds taken to perform the actions described. To help the experimenters adhere to the events' scripts, a metronome beat softly once per second. A camera mounted behind and next to the infant captured an image of the events, and a second camera mounted beneath the apparatus floor captured an image of the infant; the two images were combined and projected onto a television set located behind the apparatus, and then recorded onto a computer. During the test session, the supervisor monitored the events on the TV set to confirm that they
followed the prescribed scripts. Recorded sessions were also checked offline to ensure that the correct actions were performed in each trial.

During trials, the experimenters never made eye contact with the infant; they looked at the objects they acted on, followed each other’s actions, or (when no action occurred) looked at a neutral point on the apparatus floor.

When a trial ended, a supervisor lowered the curtain at the front of the apparatus, and stimuli were readied for the next trial. Inter-trial intervals lasted about 15 s, and each new trial began with the raising of the curtain.

**Familiarization trial**

The familiarization trial served to familiarize the infant with A1’s block building activities, and it lasted 31 s. At the start of the trial, A1 sat at the right window and held the basket of alphabet blocks. After a pause (3 s), A1 placed the basket on the apparatus floor (located 2 cm from the back wall and 12 cm from the right wall) and withdrew her hands to the window ledge (3 s). As B1 and the target individual looked on, A1 then stacked the four D-blocks into a tower with the letter Ds facing the child (12 s). Then A1 stacked the four C-blocks into a tower with the letter Cs again facing the child (12 s). After A1 reached into the basket for another block (1 s), the computer signaled the end of the trial and the supervisor lowered the curtain in front of the apparatus.

**Test trial**

The test trial consisted of an initial and a final phase. The initial phase lasted 27 s. B1 was absent (the left window was closed with an opaque white curtain). After a pause (3 s), the target individual started to eat crackers from a plate in front of her. At the same time, A1 stacked the four F-blocks into a tower with the letter Fs facing the child, looking at the target individual and her crackers after each block was placed on the tower (12 s). After A1 placed the final block on the tower, she paused to admire the three towers (2 s).

In the *provocation situation*, A1 then looked at the target individual’s crackers (2 s) and grabbed a cracker from the plate (1 s). The target individual then quickly covered her plate with both hands as A1
quickly stuffed the cracker into her mouth and started chewing it (1 s). Next, A1 left the apparatus by closing the side curtain (2 s). The no-provocation situation was similar except that, after A1 finished building her last tower, she left without taking the target individual’s cracker. Instead, A1 turned to the target individual and said, “Oh, I have to go. I will be back!” (4 s) before leaving by closing the curtain (2 s).

After A1 was gone, the target individual looked at A1’s towers (1 s). In the one-block event, the target individual knocked down the top block of the D-tower by a single push with her left hand; she then withdrew both hands on her laps, looked down at her crackers, and paused (3 s). In the some-towers event, the target individual knocked down the one tower (the D-tower) or two towers (the D- and the C-tower) by a single push with her left hand and then withdrew her hands, looked down at her crackers, and paused (3 s). During the final phase, the infants watched the final paused scene until the computer signaled the end of the trial, and the supervisor lowered the curtain in front of the apparatus.

Procedure

Each infant sat on a parent’s lap centered about 45 cm in front of the apparatus. Parents were instructed to remain silent and neutral and to close their eyes during the test trial. Prior to the session, A and B briefly introduced themselves to the infant.

Two naive observers monitored the infant’s looking behavior through peepholes in the cloth-covered frames on either side of the apparatus. Each observer depressed a button linked to a computer when the infant looked at the events shown during a trial. Looking times during the initial and final phases of each trial were computed separately, using the primary observer’s responses. Interobserver agreement during the final phases of test trials was calculated for 25/28 infants (only one observer was present for the other infants) and averaged 98% per trial per infant.

The infants were highly attentive during familiarization trial and looked for 30.9/31 s on average.

Next, each infant watched a single test trial; half the infants saw the one-block event, and half saw the some-tower event. The infants were attentive and looked for 26.4/27 s on average during the initial phase. The final phase ended when the infant (1) looked away for 0.5 consecutive seconds after
having looked for at least 10 cumulative seconds, or (2) looked for 60 cumulative seconds without looking away for 0.5 consecutive seconds.

Preliminary analyses of infants’ looking times during the final phase of the test trial revealed no interaction of sex with situation (provocation or no-provocation) and event (one-block or some-towers), $F_{s} < 1$; the data was therefore collapsed across sex in the final analyses.

**Results and Discussion**

Infants’ looking times (see Figure 2) during the final phase of the test trial were compared by a $2 \times 2$ analysis of variance (ANOVA) with situation (with or without provocation) and event (one-block or some-towers) as between-subjects factors. The main effect of situation was significant, $F(1, 24) = 8.12, p < .01, \eta_{p}^{2} = 0.25$. The main effect of event was not significant, $F(1, 24) = 2.95, p = .10, \eta_{p}^{2} = 0.11$. The interaction between situation and event was not significant, $F(1, 24) = 2.87, p = .10, \eta_{p}^{2} = 0.11$. Planned comparisons indicated that in the *provocation* situation, the infants who saw the some-towers event ($M = 29.6, SD = 3.3$) looked longer than those who saw the one-block event ($M = 20.6, SD = 8.4$), $F(1, 24) = 5.36, p < .05$, Cohen’s $d = 1.41$; however, in the *no-provocation* situation, the infants who saw the one-block event ($M = 32.6, SD = 6.2$) and those who saw the some-towers event ($M = 32.6, SD = 7.3$) looked about equally long, $F(1, 24) < 1$. In addition, the planned contrast indicated that the looking times at the one-block event in the provocation situation were significantly different from those at the some-tower event in the provocation situation and those at the one-block and some-tower events in the no-provocation situation, $F(1, 24) = 19.57, p < .05$.

These results suggested that the infants viewed it as permissible to retaliate against an ingroup member with a mild negative act (i.e., knocking down a block) following provocation; however, more severe retaliation (i.e., knocking down some more towers) was unacceptable. Infants also viewed any unprovoked aggression on an ingroup member (no matter whether it was the one-block event or the some-towers event) as unacceptable.

Why do infants view it as permissible to retaliate against *ingroup* members? While it may seem contradictory to general group well-being, ingroup retaliation may be a strategy to promote cooperation...
within groups (e.g. Boyd et al., 2003; Brock & Parker, 1995; Jesen, Call, & Tomasello, 2007). Without retaliation against non-cooperators, the welfare of the entire group may be compromised. Retaliation may decrease the likelihood of recurring harmful behaviors.

Experiment 2

Experiment 2 was designed to test whether infants would expect a retaliation against an outgroup member, as in Experiment 1; we compared the magnitude of the negative action expected with versus without provocation. Whether infants expect a more severe retaliation against an outgroup member than against an ingroup member was also examined. The infants were presented with similar events except that the target individual wore the same outfit as B1, marking that she belonged to B1’s group, rather than to A1’s group (see Figure 3). We predicted that infants would expect the target individual to (1) act more negatively toward an outgroup member when provoked than when unprovoked, and (2) retaliate more severely against an outgroup member compared to an ingroup member (in Experiment 1).

Method

Participants

Our design called for 32 participants. Thus far, 29 healthy term infants (16 boys and 13 girls) have participated in this experiment (17 months, 7 day to 19 months, 26 days, \(M = 18\) months, 21 days). Another 11 infants were tested but excluded, three because they were fussy (2) or distracted (1), three because the difference in their looking times during the test trial was over 2.5 standard deviations from the mean of their condition, and five because they looked for maximum allowed in test. Half of the infants were randomly assigned to the provocation (\(M = 19\) months, 2 days) and half to the no-provocation (\(M = 18\) months, 7 days) situation. In each situation, half the infants saw the one-block event, and half saw the some-towers event.
The apparatus and stimuli were the same as those in Experiment 1, except that the target individual and B1 both wore the bear-style outfit, while A1 wore the giraffe-style outfit, signaling that the target individual and A1 belonged to different groups.

Trials

The familiarization and test trials were also the same as in Experiment 1 except for the aforementioned outfit changes.

Procedure

The procedure was similar to the one used in Experiment 1. The infants were highly attentive and looked for 30.7/31 s on average during the familiarization trial and for 26.9/27 s during the initial phase of the test trial. Interobserver agreement during the final phase of the test trial was calculated for 28/29 infants (only one observer was present for the other infants) and averaged 99% per infant.

Preliminary analyses of infants’ looking times during the final phase of the test trial revealed no interaction of sex with situation (provocation or no-provocation) and event (one-block or some-towers), $F < 1$; the data was therefore collapsed across sex in the final analyses.

Results and Discussion

Infants’ looking times (see Figure 4) during the final phase of the test trial were compared by an ANOVA with situation (with or without provocation) and event (one-block or some-towers) as between-subjects factors. The interaction between situation and event was significant, $F (1, 25) = 19.8, p < .001$, $\eta_p^2 = 0.44$. Neither situation nor event had a significant main effect, $F < 1$. Planned comparisons indicated that in the no-provocation situation, the infants who saw the some-tower event ($M = 35.0, SD = 11.1$) looked reliably longer than those who saw the one-block event ($M = 20.1, SD = 5.6$), $F(1, 25) = 8.42, p < .05$, Cohen’s $d = 1.51$; in the provocation situation, the infants who saw the one-block event ($M = 36.2, SD = 13.9$) looked reliably longer than those who saw the some-towers event ($M = 19.3, SD = 5.6$), $F(1, 25) = 11.56, p < .05$, Cohen’s $d = 1.59$. 
These results suggested that the infants viewed it as permissible to inflict a mild unprovoked aggression on an outgroup member (i.e., knocking down a block), while more severe aggression (i.e., knocking down some more towers), was unacceptable. However, following provocation, infants viewed it as acceptable to retaliate against an outgroup member with a moderate negative act (i.e., knocking down some more towers); but a mild retaliation (i.e., knocking down only one block) was viewed as not enough.

The infants’ looking times in Experiments 1 and 2 were analyzed by an ANOVA, with condition (ingroup or outgroup), situation (with or without provocation), and event (one-block or some-towers) as three between-subject factors. The 3-way interaction was significant, \( F(1, 49) = 20.68, p < .001, \eta_p^2 = 0.30 \). Planned comparisons indicated that (1) in the no-provocation situation, the infants who saw the one-block event in the ingroup condition looked reliably longer than those who saw the same event in the outgroup condition, \( F(1, 40) = 8.35, p < .05, \eta_p^2 = 0.17 \), however, infants’ looking times at the some-towers event were approximately the same in the ingroup and outgroup conditions, \( F(1, 40) < 1 \); (2) in the provocation situation, the infants who saw the one-block event in the outgroup condition looked reliably longer than those who saw the same event in the ingroup condition, \( F(1, 40) = 13.01, p < .05, \eta_p^2 = 0.25 \), however, the infants who saw the some-towers event in the ingroup condition looked reliably longer than those who saw the same event in the outgroup condition, \( F(1, 40) = 4.67, p < .05, \eta_p^2 = 0.10 \).

These results suggests that infants expect a more severe retaliation in the outgroup condition than in the ingroup condition: a mild retaliation is expected in the ingroup condition, whereas a moderate retaliation is expected and a mild retaliation is not enough in the outgroup condition.

Why do infants expect individuals to retaliate against outgroup members more severely than against ingroup members? The norm of ingroup loyalty that individuals should act in ways to support their groups may dampen the retaliation against ingroup members. The main purpose of punishing non-cooperative members is to make those group members more cooperative but not to harm them. Therefore, a relative escalation of retaliation is expected in the outgroup compared to the ingroup condition.
Experiment 3

Experiment 3 examined whether infants expect any limit on the severity of retaliation against an outgroup member. The infants were presented with the all-towers events in which the target individual knocked down all three of A1’s towers after A1 stole and ate her cracker (see Figure 5).

If the infants generally viewed knocking down all three towers as too severe a retaliation against someone who had stolen one of a plentiful plate of crackers, then the infants in the ingroup and the outgroup condition might all view the all-tower event as unexpected.

Method

Participants

Our design called for 16 participants. Thus far, 12 healthy term infants, 7 male and 5 female (17 months, 26 day to 19 months, 25 days, $M = 18$ months, 22 days) have been tested. Another infant was tested but excluded because she looked for the maximum allowed in the test trial. Half of the infants were randomly assigned to the ingroup and half to the outgroup condition.

Apparatus and Stimuli

The apparatus and stimuli were the same as those used in Experiments 1 and 2.

Trials

Familiarization trial

Infants in the ingroup condition saw the same familiarization trial as in Experiment 1, while those in the outgroup condition saw the same one as in Experiment 2.

Test trial

The test trial was similar to the one in the provocation situation of Experiments 1 and 2 except that the target individual knocked down all the three towers by a single push with left hand (all-towers event).
Procedure

The procedure was the same as in Experiments 1 and 2. The infants were highly attentive and looked for 30.6/31 s on average during the familiarization trial and for 27/27 s during the initial phase of the test trial. Interobserver agreement during the final phase of the test trial was calculated for 11/12 infants (only one observer was present for the other infant) and averaged 94% per infant.

Preliminary analyses of infants’ looking times during the final phase of the test trial revealed no interaction of sex with condition (ingroup or outgroup), $F(1,8) < 1$; the data was therefore collapsed across sex in the final analyses.

Results and Discussion

Infants’ looking times during the final phase of the test trial were compared by a $t$-test with condition (ingroup or outgroup) as a between-subjects factor. The children in the ingroup ($M = 34.8, SD = 13.6$) and in the outgroup ($M = 34.1, SD = 14.5$) condition looked at the all-towers event about equally, $t(10) < 1$.

In additional analyses, infants’ looking times at the all-towers events were contrasted to those of the infants in the provocation condition in Experiments 1 and 2 (see Figure 6). The infants’ looking times were analyzed by an ANOVA, with condition (ingroup or outgroup) and event (one-block, some-towers, or all-towers) as two between-subject factors. The interaction between condition and event was significant, $F(2, 34) = 5.15, p < .05, \eta^2_p = 0.23$. We first compared the looking times at the one-block, some-towers, and all-towers events of the infants in the ingroup condition. The planned contrast showed that the children who saw the all-towers and some-towers events looked reliably longer than those who saw the one-block event, $F(1, 34) = 5.42, p < .05, \eta^2_p = 0.14$. These results suggest that infants expect mild to moderate retaliation against an outgroup member, and that detect a violation if the retaliation become moderate to severe. Next, we compared the looking times at the one-block, some-towers, and all-towers events of the infants in the outgroup condition. The planned contrast showed that the children who saw the one-block event and the all-towers event looked reliably longer than those who saw the some-towers event, $F(1, 34) = 16.76, p < .05, \eta^2_p = 0.33$. These results suggest that infants expect a moderately-
severe level of retaliation against an outgroup member, and that detect a violation if the retaliation is either too mild or too severe.

Together, these findings suggest that infants expect more a severe retaliation against outgroup members, but that there is still a limit to the severity of the retaliation that is permissible against them. What limits the retaliation against outgroup members? There may be several possibilities. (1) If an individual retaliates against an outgroup member too severely, the overreaction may lead to a new cycle of retaliation by that outgroup member; (2) The welfare and fairness principles, which apply across groups, may contain retaliations across groups. For example, the principle of “do no harm” may limit how severely individuals act negatively toward outgroup members (e.g. Premack, 2008). Moreover, the principle of fairness may require the severity of punishment for a misdeed or wrongdoing to be reasonable and proportional to the severity of the infraction, regardless of whether the target of the punishment is an ingroup or an outgroup member.

**Discussion of Project 1**

Overall, the results of the three experiments in Project 1 suggest that 19-month-old infants (1) expect individuals to retaliate against both ingroup and outgroup members and (2) expect a more severe retaliation against outgroup than against ingroup members.

As far as we know, Project 1 is the first to examine infants’ understanding of the magnitudes of actions in addition to their valence. The results suggest that infants, by 19 months of age, are already sensitive to the magnitude of retaliatory actions, and form different expectations when the two individuals belong to the same group or different groups. Future investigation is needed to examine whether the difference in magnitude of retaliation expected against outgroup versus ingroup members were due to amplified retaliation in the outgroup condition (related to an unmarked condition in which group membership is not indicated) or if it is due to dampened retaliation in the ingroup condition (related to the unmarked condition).

The results of Project 1 also provide evidence on infants’ expectations about the valence of individuals’ initial (as opposed to reciprocal) actions. In our experiments, infants consider it permissible
to inflict a mild unprovoked aggression on an outgroup member, but unacceptable on an ingroup member. These findings are consistent with prior evidence on the early sensitivity of ingroup loyalty (e.g., Kinzler et al., 2007, 2011; Mahajan & Wynn, 2010; Over & Carpenter, 2009; Shutts et al., 2009; Werneken & Tomasello, 2006, 2007).

How do infants attain their expectations about retaliation? One possibility is that reciprocity is one of a small set of innate norms guiding socio-moral reasoning (e.g., Dwyer, 2007; Haidt & Joseph, 2007; Premack, 2007). Another possibility is that infants identify retaliation as one of the norms that prevail in their daily social environment, and gradually extend it to new individuals and situations (e.g., Sripada & Stich, 2006; Turiel, 2006). In a preliminary analysis combining the data from the three experiments, we found no reliable difference in the test responses of infants with and without older siblings, nor did we find a reliable difference in the test responses of infants with and without substantial day care experience. This suggests that interactions with peers may not be necessary for an expectation of retaliation, and reciprocity more generally, to emerge.

These early expectations about aggression and retaliation must help infants interpret the social interactions they observe daily and predict the consequences of their own actions toward others. These expectations may lay the foundation for the children to understand and deal with social conflicts in their relations with others. It is important to investigate in the future the effect of social environments on these early expectations, for example, whether and how these expectations would be modified when children live in an abusive, as opposed to a supportive, environment.

In sum, the findings of Project 1 indicates that, by the second year of life, infants already possess expectations about retaliation and these expectations are modulated by considerations of ingroup loyalty. Infants expect an escalation of retaliation in the outgroup condition compared to the ingroup condition. That is, infants expect a more severe retaliation against an outgroup member than against an ingroup member.
CHAPTER 3: PROJECT 2 ON CO-RETAIATION

Project 1 focused on infants’ expectations about the interactions between A1 and the target individual (B1 was present in the familiarization trial only to make the group markers more salient). The results suggest that infants expect the target individual to act negatively in response to A1’s provocation, with a more severe magnitude when A1 is an outgroup member. This is an example of first-party punishment in which victims directly punish those who are responsible for causing them harm.

One question that was not addressed is whether infants also hold expectations about third-party punishment. This occurs when punishment is enacted by someone who was not directly harmed by the actions of the individual who is being punished. Third-party punishment is critical for preventing social norms from being violated and for maintaining cooperation and positivity within a group. In human history, complex social organizations, such as government and police, have evolved so that some people are elected to take the responsibility of punishing group members who violate the interest of the group (Fry, 2006). Prior research on third-party punishment suggests that people are willing to punish non-cooperators in their groups even when they themselves do not directly benefit from the punishment (e.g., Fehr, & Fischbacher, 2004; Gummerum et al., 2009).

The idea of co-retaliation as proposed by David Premack expands on this idea of third-party retaliation by including the factors of ingroup and outgroup: Assuming B1 and B2 belong to the same group, and A1 belongs to a different group, if A1 hinders B1, then B2 should retaliate against A1 on behalf of B1 (Premack & Premack, 1995).

Prior research suggests that belonging to the same group as the victim or the perpetrator of a moral transgression affects adults’ decision about the punishment assigned to the perpetrator. In one study, participants read scenarios describing a moral violation in which the victim of an offense was described as kin, a schoolmate, or a foreigner (Lieberman & Linke, 2007). Here is an example of one of the scenarios: a burglar broke into your family member’s home and stole some expensive property. Results showed that the participants wanted the burglar to serve the longest jail time or to pay the highest fine if the victim was their family member, followed by a schoolmate, and then a foreigner. The participants also reported a
greater willingness to sacrifice their time or money to punish a perpetrator who victimized their family member, again followed by a schoolmate, and then a foreigner. In addition, the allocated punishment was affected by whether the perpetrator was an ingroup or an outgroup member; family members received the least punishment and outgroup members received the greatest punishment.

This co-retaliation may be an important manifestation of ingroup loyalty: people provide assistance and protection to ingroup members, and they co-retaliate against outgroup individuals who have acted negatively toward ingroup members. Understanding the norm of ingroup loyalty is critical for one’s survival: as we know from human history, people can be executed for treason after helping the enemies of their groups. This norm is also found in groups of non-human primates. For example, studies on reciprocal grooming behavior among chimpanzees found that chimpanzees tend to concentrate benevolent grooming behavior only on their core alliances. Members outside of this core group were not treated equally (e.g. Watts, 2000). This indicates that group loyalty requires individuals to direct more prosocial behaviors to ingroup than outgroup members.

Project 2 was aimed at exploring how children would respond if an individual helped or hindered an outgroup member who had previously hindered the first individual’s ingroup member (i.e., co-retaliation). Such exploration could help expand the current understanding of children’s expectations of the interactions among ingroup and outgroup individuals. The following questions were addressed. (1) What are children’s expectations about triadic interactions of individuals (as opposed to the dyadic interactions explored in Project 1)? (2) Do children keep track of prior interactions among individuals, and consider the history of their interactions when predicting and interpreting subsequent interactions among the individuals? (3) Do children expect third-party retaliation against an outgroup member who has acted negatively against an ingroup member? (4) Do children detect a violation when individuals betray their ingroup, by helping an outgroup member who has hindered an ingroup member?

In Experiments 4 and 5, we examined whether infants would expect an individual to retaliate against an outgroup member on behalf of her ingroup member (i.e., the expectation of co-retaliation) when the group membership was marked by contrasting outfits (as in Experiments 1 through 3).
Experiments 6 and 7, we examined whether toddlers would show the expectation of co-retaliation, when the group membership was marked only by novel naming labels. We reasoned that positive findings in both infants and toddlers would provide converging evidence for the expectation of co-retaliation and would also suggest the continuity of early socio-moral expectations prior to preschool years.

**Experiment 4**

The findings of Experiments 1 through 3 and a prior study (e.g., He et al., under review) suggest that if a first individual acts negatively toward another individual, infants expect the second individual to act negatively in response to the provocation. Here, we examined whether infants’ expectations of retaliation would be modulated by considerations of ingroup loyalty. Experiment 4 included 3 conditions that depicted events involving members of two arbitrary social groups, A and B. Children first saw A1 act negatively toward B1. We asked whether infants would expect B2, a member of the same group as B1, to react negatively toward A1. Positive results would suggest that ingroup loyalty entails co-retaliation against outgroup individuals who have acted negatively toward ingroup members.

In the *outgroup* condition, infants saw A1 act negatively towards B1. Then the target individual, also a member of group B, was given an opportunity to either help or hinder A1 (see Figure 7). Infants watched four familiarization trials, one pretest trial, and one test trial. At the start of each familiarization trial, three individuals sat at an apparatus. A1, on the right, and B1, on the left, wore different outfits and thus belonged to different groups; the target individual (who acted either positively or negatively toward A1 in the test trial) sat in the middle and wore the same outfit as B1. In the first two familiarization trials, as A1 and the target individual looked on, B1 made a rattle by placing a marble inside a plastic egg and shaking it; in the next two trials, A1 stole B1’s marble and left. In test, B1 was absent, and A1 stacked rings on a pole to make a tower. After vainly attempting to reach for the final ring, which lay across the apparatus out of reach, A1 left. B2 then either threw A1’s ring away presumably so that A1 could not finish her tower when she returned (*hinder* event) or made it accessible to A1 when she returned (*help* event). We predicted that the infants would expect the target individual to hinder rather than help A1, since she had previously hindered the B2’s ingroup member, B1.
Two alternative interpretations besides the expectation of co-retaliation were also taken into consideration in this experiment. The first alternative interpretation is that infants might just expect the target individual, B2, to hinder anyone who had hindered regardless of whether they were an ingroup or an outgroup member. It is possible that infant have a general bias against antisocial individuals, no matter whether they are ingroup or outgroup members. For example, in the experiments of Hamlin, Wynn, and Bloom (2007), 6- and 10-month-old infants (1) preferred prosocial (helping) individuals to antisocial (hindering) individuals, (2) preferred prosocial to neutral individuals (who neither helped nor hindered), and (3) preferred neutral to antisocial individuals. Moreover, in the follow-up study (Hamlin, Wynn & Bloom, 2010), 3-month-olds showed an aversion to antisocial individuals—they preferred neutral to antisocial individuals, though they had no particular preference between neutral and prosocial individuals.

To rule out this alternative interpretation, an *ingroup* condition was included (see Figure 8). The infants in the ingroup condition watched similar trials except that the target individual wore the same outfit as A1. We predicted that the infants would expect the target individual to help rather than hinder A1 who belonged to the same group as the target individual, despite the fact that A1 had hindered an outgroup member.

The other interpretation was that infants might just expect the target individual to hinder the outgroup member, regardless of whether they had hindered or not, due to a sort of generalized “outgroup” hatred. Prior research on social cognition in adults suggests that intergroup conflicts are in fact mostly driven by “ingroup love” rather than by “outgroup hatred”. “Outgroup hatred” is not likely a part of our human nature, but a byproduct of resource sharing and exchange that takes place within the boundaries of a group (e.g., Brewer, 1999; Yamagishi & Mifune, 2009).

To rule out this alternative interpretation, an *outgroup-no-provocation* condition was included (see Figure 9). The infants in the outgroup-no-provocation condition received trials similar to the outgroup condition except that the infants did not received the familiarization trials that showed A1 hindering B1. We predicted that the infants would not have particular expectations about whether the target individual should help or hinder the outgroup member in this no-provocation condition.
To sum up, the results of the outgroup, the ingroup, and the outgroup-no-provocation conditions together might provide evidence that infants, by 13 months, are already sensitive to novel social groups as arbitrarily marked by contrasting outfits and have an expectation of co-retaliation.

Method

Participants

Our designed called for 48 participants. Thus far, 47 healthy term infants (24 male and 23 female) have been tested (12 months, 6 day to 14 months, 1 days, \( M = 13 \) months, 9 days). Another 10 infants were tested but excluded, six because they were either fussy (2), inattentive (2), active (1), or had a looking time that was over 2.5 standard deviations from the mean of their condition in the test trial (1), and four because they looked for the maximum allowed in the test. Sixteen infants were randomly assigned to the outgroup (\( M = 13 \) months, 6 days) and ingroup (\( M = 13 \) months, 8 days) conditions, and fifteen to the outgroup-no-provocation (\( M = 13 \) months, 12 days) condition. In each condition, half the infants saw the hinder event, and half saw the help event.

Apparatus and Stimuli

The apparatus was the same as in Experiments 1 through 3.

Three female experimenters were involved in the events. A1 knelt at a curtained window in the right wall of the apparatus. A1 wore an orange hooded sweatshirt with the hood on and tied at the chin. Her hair was tied back and she also wore a pair of purple clear glasses.

B1 knelt at the left window of the apparatus and wore a black long-sleeve turtleneck. Her hair was tied back and she wore a pair of black-and-orange plush tiger-like ears and a black-and-orange plush tiger-like collar around her neck.

The target individual knelt at the back window in an outfit that depended on the condition. She wore the same outfit as B1 in the outgroup and the outgroup-no-provocation conditions, while she wore the same outfit as A1 in the ingroup condition.
At the beginning of the first two familiarization trials, the two halves of a plastic Easter egg (6 cm X 4.5 cm in diameter at its largest points) were separated and placed with their open ends down on the apparatus floor. The Easter egg was pink in the first trial and green in the second trial. One half was located 18 cm in front of the back window, and 6.5 cm from the left window; the other half was located 33 cm in front of the back window, and 6.5 cm from the left window. A white placemat (0.5 cm X 8.5 cm X 8.5 cm) was placed on the apparatus floor 22.5 cm in front of the back window, and 24 cm from the left window. Centered on the placemat was a green marble (0.5 cm X 2 cm in diameter).

The setup of the next two familiarization trials was the same as in the first two trials except that the placemat lay across the apparatus floor near A1, located 16 cm from the left window and 39 cm in front of the back window. The plastic Easter egg was yellow in the third trial and purple in the fourth trial.

At the beginning of the pretest trial, the base of a stacking ring set (19 cm X 13 cm X 13 cm at its largest points) stood on the apparatus floor, located 29.5 cm in front of the back window and 2 cm from the left window. The base had a white bottom with a yellow tapered pole in its center. A white wooden tray (4 cm X 21 cm X 16.5 cm) was placed on the apparatus floor 2 cm in front of the back window and 2 cm from the right window. The tray contained a stacking ring set that included, from largest to smallest, a blue, a green, a yellow and an orange ring. Each ring was placed against one edge of the tray.

The setup for the test trial was the same as in the pretest except that the orange ring was not in the tray, but lay across the apparatus floor, about 6 cm from the left window and 22 cm in front of the back window.

Trials

Familiarization trials

The first familiarization trial only had an initial phase, which lasted about 24 s. After a pause (2 s), B1 made a rattle by placing the marble inside the cup (11 s) and shaking it while the target individual and A1 looked on (11 s). Next, the computer signaled the end of the trial, and the supervisor lowered the curtain in front of the apparatus. The second familiarization trial was identical to the first one except that the egg was replaced with one of a different color.
The third familiarization trial lasted about 34 s. The setup was the same as in the first two trials except that the placemat lay across the apparatus floor near A1. After a pause (2 s), B1 attempted but failed to reach for the marble while the target individual and A1 looked on (7 s). Then A1 took the marble away and left the apparatus by closing the curtain of the right window of the apparatus (6 s). B1 looked down and appeared unhappy (19 s). Next, the computer signaled the end of the trial, and the supervisor lowered the curtain in front of the apparatus. The fourth familiarization trial was identical to the third one except that the egg was replaced with one of a different color.

Pretest trial

The pretest trial only had an initial phase, which lasted 21 s. B1 was absent (the left window was completely closed with an opaque white curtain). After a pause (2 s), A1 stacked the blue, the green, the yellow, and then the orange ring on the pole as the target individual looked on (16 s). She withdrew her hand and looked at the stacking ring set pleasantly (3 s). Next, the computer signaled the end of the trial, and the supervisor lowered the curtain in front of the apparatus. The pretest trial served to make clear that A1 had a goal of stacking her rings and enjoyed accomplishing the goal.

Test trial

The test trial consisted of an initial and a final phase. The setup was the same as in the pretest except that the orange ring was not in the tray, but lay across the apparatus out of reach. The initial phase lasted for 29 s. B1 was absent. After a pause (2 s), A1 stacked the blue, the green, and then the yellow ring on the pole as the target individual looked on (12 s). Then A1 vainly attempted to reach for the orange ring, which lay across the apparatus floor (6 s). A1 left the apparatus by closing the curtain in the right window of the apparatus (2 s).

In the hinder event, the target individual grasped the orange ring with her right hand (2 s) and threw the orange ring out of the apparatus (4 s). She then withdrew her hand to the apparatus floor, looked down at a neutral point, and paused (1 s).

In the help event, the target individual grasped the orange ring with her right hand (2 s) and put it into A1’s tray (4 s). She then withdrew her hand, looked down at a neutral point, and paused (1 s).
During the final phase, the infants watched the final paused scene until the computer signaled the end of the trial, and the supervisor lowered the curtain in front of the apparatus.

**Procedure**

The infants in the outgroup condition and those in the ingroup condition watched four familiarization trials, one pretest trial and one test trial. The infants in the outgroup-no-provocation condition watched only the first two familiarization trials, one pretest trial and one test trial; hence they never saw A1 hinder B1.

The infants were highly attentive and looked for 23.9/24 s on average during the first two familiarization trials, 33.5/34 during the last two familiarization trials, 20.8/21 during the pretest trial, and 28.9/29 s during the initial phase of the test trial.

The final phase of the test trial ended when the infant (1) looked away for 0.5 consecutive seconds after having looked for at least 10 cumulative seconds, or (2) looked for 40 cumulative seconds without looking away for 0.5 consecutive seconds.

Interobserver agreement during the final phase of the test trial was calculated for 42/47 infants (only one observer was present for the other infants) and averaged 99% per infant.

Preliminary analyses of infants’ looking times during the final phase of the test trial revealed no interaction of sex with condition (outgroup, ingroup, or outgroup-no-provocation) and event (help or hinder), $F(2, 35) = 2.00, p > .05$; the data was therefore collapsed across sex in the final analyses.

**Results and Discussion**

Infants’ looking times (see Figure 10) during the final phase of the test trial were compared by an ANOVA with condition (outgroup, ingroup, or outgroup-no-provocation) and event (help or hinder) as between-subjects factors. The interaction between condition and event was significant, $F(2, 41) = 9.54, p < .001, \eta^2_p = 0.32$, but the main effects of condition and event were not significant, $Fs < 1.35, ps > .05$. The planned comparisons revealed that children in the outgroup condition who saw the help event ($M = 23.6, SD = 8.2$) looked reliably longer than those who saw the hinder event ($M = 15.1, SD = 4.0$), $F(1, 41) = 7.71, p < .05$, Cohen’s $d = 1.32$; children in the ingroup condition showed the reversed pattern: the
children who saw the hinder event ($M = 24.7, SD = 8.7$) looked reliably longer than those who saw the help event ($M = 14.3, SD = 3.0$), $F(1, 41) = 11.54, p < .05$, Cohen’s $d = 1.60$; in the outgroup-no-provocation condition, the children who saw the hinder event ($M = 16.6, SD = 5.6$) and those who saw the help event ($M = 15.9, SD = 5.0$) looked about equally long, $F(1, 41) < 1, p > .05$.

In the outgroup condition, infants who saw the help event looked reliably longer than those who saw the hinder event. Thus, when $A1$ hindered $B1$ and then needed assistance to achieve her own goal, infants expected $B2$ to hinder $A1$, in co-retaliation for $A1$’s prior negative actions toward $B1$.

By comparing the results of the outgroup condition to the ingroup condition and the outgroup-no-provocation condition, we can rule out the two alternative interpretations. Infants in the outgroup condition expected the target individual to hinder $A1$, because they expected the target individual to hinder any hinderer, nor because they expected the target individual to hinder any outgroup member. The infants in the ingroup condition expected the target individual to help $A1$, irrespective of $A1$’s prior negative actions toward an outgroup member. This expectation of assistance is consistent with the evidence that young children are more likely to provide assistance to an ingroup member or to a stranger when group affiliation is primed (e.g., Over & Carpenter, 2009; Werneken & Tomasello, 2006, 2007).

The infants in the outgroup-no-provocation condition had no particular expectation about whether the target individual should help or hinder $A1$ without provocation, suggesting that the infants clearly expected the target individual to hinder $A1$ only after $A1$ first hindered $B1$. This finding provides evidence supporting the claim that ingroup bias is rooted in “ingroup love” rather than in “outgroup hatred” (e.g., Brewer, 1999; Yamagishi & Mifune, 2009).

The results of the three conditions together suggest that infants attend to group membership and expect individuals (1) to co-retaliate against outgroup individuals who have acted negatively toward ingroup members, but (2) to help ingroup members in need of assistance irrespective of whether they have acted negatively toward outgroup individuals.

Experiment 4 provides evidence for early expectations of co-retaliation: infants expect an individual to retaliate against an outgroup individual on behalf of a group member. However, another two
alternative interpretations have been offered. First, expectations of co-retaliation could be attributed to infants’ own emotional responses to the provocation. If the infants sympathized with the victim of the provocation, they could have become upset and thus looked longer when they saw someone help the perpetrator of the provocation (sympathy hypothesis).

Another alternative interpretation is that infants represent group members solely as tokens of the group, and that the tokens are identical and interchangeable. In other words, infants either are simply confused between different group members or they assume all group members share the same mental states (token hypothesis). Therefore, infants expected any group member to co-retaliate when any group member was provoked, and thus looked longer when they saw a group member help the perpetrator of the provocation.

However, prior evidence suggests that infants are more capable than what the sympathy or the token hypothesis suggests. In particular, infants can attribute different mental states to different individuals (e.g., Baillargeon et al., 2010; Johnson, 2000; Luo & Baillargeon, 2011). Therefore, it seemed more likely that infants could represent both the group memberships and the distinct mental states of the individuals in the events (mentalistic hypothesis). The mentalistic hypothesis argues that infants can attribute differentiated mental states to each member of the group; they can make sense of intentional actions in terms of individuals’ goals, knowledge, and other mental states, even when those mental states are inconsistent with infants’ own mental states. Therefore, only the group members who know about the provocation targeted at another group member would be expected to co-retaliate.

**Experiment 5**

In the previous experiment, we hypothesized that infants would expect an individual to retaliate against an outgroup individual on behalf of another ingroup member. We reasoned that this behavior would be modulated by infants’ ingroup loyalty. While our hypothesis was confirmed, two alternative explanations do exist. As detailed in the discussion section of Experiment 4, infants could be influenced by their emotional involvement in the scene that they witnessed (sympathy hypothesis), or they could be reacting to group members as identical and interchangeable tokens (token hypothesis). We believe that a
third hypothesis, the mentalistic hypothesis compliments the reasoning we used in Experiment 4. Experiment 5 was designed to test the mentalistic hypothesis, while ruling out the sympathy and token hypotheses.

Experiment 5 manipulated whether the provocation occurred during the target individual’s absence or presence. First, this manipulation would control for children’s emotional involvement in the scene. Children’s own emotion reactions should be equivalent whether the target individual was absent or present. Hence, the sympathy hypothesis predicted that the infants would expect the target individual to co-retaliate regardless of whether she was ignorant or knowledgeable about the provocation.

Second, this manipulation would also examine whether children represented group member solely as interchangeable tokens. This token hypothesis argued that infants were simply confused between different group members or they assumed all group members share the same mental states. Hence, the token hypothesis, similar to the sympathy hypothesis, predicted that infants would expect any group member to co-retaliate regardless of whether she was ignorant or knowledgeable about the provocation.

Therefore, three hypotheses were tested in Experiment 5. We asked whether the target individual’s knowledge about the provocation would affect infants’ expectation of co-retaliation. We proposed that infants would expect the target individual to co-retaliate only when she knew about the provocation.

In the outgroup condition (see Figure 11), A1 acted negatively towards B1 in the absence or in the presence of the target individual, also a member of group B. Then the target individual helped A1. (Given the complementary nature of helping and hindering behaviors, we only included the help event in this experiment.) We predicted that (1) if the target individual was present during a provocation targeted at her ingroup member, then infants would consider it unacceptable for the target individual to help the outgroup perpetrator; but (2) if the target individual was absent during the provocation, then infants would consider it acceptable for the target individual to help the outgroup perpetrator. The first results would confirm the findings from the outgroup condition of Experiment 4 that infants expect individuals to co-retaliate against outgroup individuals who have acted negatively toward ingroup members; whereas
the second results would confirm the findings from the outgroup-no-provocation condition of Experiment 4 that infants have no particular expectations whether individuals should help or hinder outgroup members without provocation.

In the ingroup condition (see Figure 12), the target individual was portrayed as an ingroup member of perpetrator A1. We predicted that infants would consider it acceptable for the target individual to help her ingroup member A1 regardless of whether the target individual was present or absent when A1 hindered the outgroup member, B1.

In other words, we predicted that in the outgroup condition, infants who saw the help event in the presence situation should look reliably longer than those who saw the help event in the absence situation. In the ingroup condition, infants who saw the help in the presence situation and those who saw the help event in the absence situation would look about equally long.

Method

Participants

Our designed called for 32 participants. Thus far, 31 healthy term infants (16 male and 15 female) have been tested (13 months, 2 day to 14 months, 12 days, $M = 13$ months, 22 days). Another 6 infants were tested but excluded, three because they were fussy (1) or had a looking time that was over 2.5 standard deviations from the mean of their condition in the test trial (2), and three because they looked for the maximum allowed in the test trial. Eight infants were randomly assigned to the outgroup ($M = 13$ months, 21 days) and the ingroup ($M = 13$ months, 22 days) conditions. In each condition, half the infants saw that the target individual leave the apparatus before the provocation occurred (absence situation) and half saw that she leave after the provocation occurred (presence situation).

Apparatus and Stimuli

The apparatus and stimuli were the same as in Experiment 4.
Trials

Familiarization trials

The first two familiarization trials were identical to those used in Experiment 4.

The last two familiarization trials were similar to those in Experiment 4 except that the target individual left either before or after the provocation (see the detailed descriptions of the actions below). The initial phase lasted 38 s for both condition.

In the absence situation, after a pause (2 s), the target individual left the apparatus by closing the doors of the back window (4 s). As A1 looked on, B1 vainly attempted to reach for the marble (7 s). A1 took the marble and left the apparatus by closing the right window curtain of the apparatus (6 s). B1 looked down and appeared unhappy (19 s).

In the presence situation, after a pause (2 s), B1 attempted but failed to reach for the marble while the target individual and A1 looked on (7 s). A1 took the marble and left the apparatus by closing the right window curtain of the apparatus (6 s). B1 looked down and appeared unhappy (19 s). Then the target individual left the apparatus by closing the doors of the back window (4 s).

Pretest trial

The pretest trial was identical to that in Experiment 4.

Test trial

The test trial was identical to the help event test trial in Experiment 4.

Procedure

The procedure was the same as in Experiment 4.

The infants were highly attentive and looked for 23.9/24 s on average during the first two familiarization trials, 37.1/38 during the last two familiarization trials, 20.9/21 during the pretest trial and 28.9/29 s during the initial phase of the test trial.

The final phase of the test trial ended when the infant (1) looked away for 0.5 consecutive seconds after having looked for at least 10 cumulative seconds, or (2) looked for 40 cumulative seconds without looking away for 0.5 consecutive seconds.
Interobserver agreement during the final phase of the test trial was calculated for 23/31 infants (only one observer was present for the other infants) and averaged 97% per infant.

Preliminary analyses of infants’ looking times during the final phase of the test trial revealed no interaction of sex with condition (outgroup or ingroup) and situation (presence or absence), $F(1, 23) = 1.12, p > .05$; the data was therefore collapsed across sex in the final analyses.

**Results and Discussion**

Infants’ looking times (see Figure 13) during the final phase of the test trial were compared by an ANOVA with condition (outgroup or ingroup) and situation (presence or absence) as between-subjects factors. The interaction between condition and situation was significant, $F(1, 27) = 4.54, p < .05, \eta_p^2 = 0.14$; the main effect of situation was also significant, $F(1, 27) = 6.04, p < .05, \eta_p^2 = 0.18$; the main effect of condition was not significant, $F(1, 27) = 2.13, p > .05, \eta_p^2 = 0.07$. Planned comparisons revealed that in the *outgroup* condition, the infants who saw the help event in the *presence* situation ($M = 25.7, SD = 6.8$) looked reliably longer than those who saw the help event in the *absence* situation ($M = 15.6, SD = 5.0$), $F(1, 27) = 10.1, p < .05$, Cohen’s $d = 1.69$. The looking time of the infants in the *ingroup* condition who saw the help event in the *presence* situation ($M = 17.8, SD = 6.4$) and those who saw the help event in the *absence* situation ($M = 17.0, SD = 6.0$) were not significantly different, $F(1, 27) < 1$.

Therefore, the infants expected the target individual to co-retaliate against the outgroup individual on behalf of her ingroup member when the target individual *herself* had witnessed the provocation but not when she was absent during the provocation. The infants expected the target individual to help her ingroup member irrespective of whether she had witnessed the ingroup member act negatively toward the outgroup individual.

These results together provide evidence for the *mentalistic hypothesis*. That is, infants can attribute differentiated mental states to each member of the group, and hence expect only those who have witnessed the provocation on other group members to retaliate against the perpetrator.

These findings of Experiment 5 rule out the token hypothesis. Infants’ expectation of co-retaliation is not due to infants’ confusion of members of a group. Infants can take into account the
knowledge or ignorance of each group member when they generate expectations about those members’ actions. Therefore, infants do not represent members of a group as identical or interchangeable tokens of the group.

These findings also rule out the sympathy hypothesis. Infants’ expectation of co-retaliation is not due to infants’ sympathy with the victim of the provocation: had this been the case, infants’ emotional response to the provocation would have consistent whether or not the target individual was present during the provocation. However, infants expected the target individual to co-retaliate against the outgroup perpetrator only if she had seen the provocation.

To sum up, the results of experiments 4 and 5 suggest that by 13 months, infants already attend to individuals’ group membership, hold different expectations for ingroup and outgroup members, and expect group members to protect one another through co-retaliation.

**Experiment 6**

Experiment 6 was aimed at examining whether 2.5-year-old toddlers would also have an expectation of co-retaliation. We reasoned that positive findings in both infants and toddlers would provide converging evidence for the expectation of co-retaliation and would also support the continuity of early socio-moral expectations in the first three years of life.

We also wanted to explore what young children would consider to be sufficient group markers. Specifically, we wanted to test whether young children would form social groups based on novel verbal labels.

Prior research shows that labeling plays an important role in children’s object categorization. For example, in a modified novelty-preference paradigm (Waxman & Markow, 1995), 12- to 13-month-old children formed a cat category after hearing the experimenter referred to different toy cats with the same label (e.g. “Look, a cat!”). This labeling effect persisted even when the experimenter used a less familiar noun, such as “See, a feline.”

Social labels may also play a role in preschoolers’ social categorization. For example, in a series of studies by Bigler and her colleague (1995, 1997, 2006), children as young as 3 years showed an
ingroup bias when groups were formed based on the color of their shirts. The children were asked to each wear either a blue or a red T-shirt. For three weeks, the teachers routinely used the group labels (i.e., the “blue team” or the “red team”) when asking children to perform tasks. By the end of the three weeks, the children showed consistent bias favoring their own group.

Gordon Allport’s *Nature of Prejudice* refers to some labels of social categories as “exceedingly salient and powerful. They tend to prevent alternative classification, or even cross-classification.” Some social labels “act like shrieking sirens, deafening us to all finer discriminations that we might otherwise perceive.” (Allport, 1954, p.179). The extraordinary power of our labels for social categories may be due to the fact that we tend to treat social categories as natural kinds rather than artificial kinds. Although social categories are in reality more like artifacts than natural kinds, social categories are often perceived as more like natural kinds than artifacts. This common misconception is probably because social categories are often believed to possess underlying essences that make one social category different from another (e.g. Haslam et al., 2000; Rothbart & Taylor, 1992).

Studies have also shown the effect of social labels on infants’ social categorization. For example, in an unpublished experiment by Waxman, infants were first presented with an example of a novel social category – all the exemplars were labeled with a nonsense noun (e.g., “a gorp”). Then the infants were introduced to a feature of the exemplar, which was also labeled with the nonsense word (e.g., He likes eating “blicket”). The results show that infants generalized the feature of the exemplar to other members of the social category (e.g., the infants expected that other “gorps” also liked eating “blicket”). These findings suggested children can form social categories based on novel naming labels. Also, the labeling effect may vary when using different grammatical forms of words. For example, children by 5 years of age expect a feature of a social category to be stable and enduring when the feature is referred to as a noun, such as “She is a carrot-eater,” but not when it is referred to as a verbal predicate, such as “She eats carrots whenever she can” (Gelman & Heyman, 1999).

In Experiment 6, we examined whether toddlers would be sensitive to social groups marked by novel naming labels. We used the novel naming labels “I am a boga!” and “I am a maloo!” as artificial
group markers. Furthermore, we examined whether toddlers would expect individuals with the same label to protect each other through co-retaliation.

Toddlers in the *outgroup* condition first watch two familiarization events (see Figures 14 & 15). At the start of each familiarization trial, three individuals sat at an apparatus: A1, on the right, B1, on the left, and the target individual. The target individual and B1 belonged to the same group while A1 belonged to another group. Group membership was identified by naming labels. A1 first announced, “I am a jaybo!” Then the target individual said, “I am a topid!” Then B1 said, “I am topid, too!” The labels were repeated twice, with order counterbalanced across toddlers. In the first familiarization trial, after the labeling, B1 made a rattle by placing a block inside a plastic cup and shaking it, as A1 and the target individual looked on. In the second familiarization event, after the labeling, A1 stole B1’s block, so that B1 could no longer make her rattle.

The toddlers then watched a single test trial: B1 was absent; A1 now was working on a puzzle as the target individual looked on. All the pieces except one were in a tray in front of her. The last piece lay across the apparatus floor, out of her reach (but within reach of the target individual). She put all the pieces on the tray in a puzzle board. She then attempted but failed to reach for the last puzzle piece. Next, A1 left and closed the curtain behind her. After A1 was gone, the target individual either threw the last puzzle piece out of the apparatus presumably so that A1 could not finish her puzzle when she returned (*hinder* event), or she put it in A1’s tray (*help* event).

As in Experiment 4, an *Ingroup* condition (see Figures 16 & 17) and an *outgroup-no-provocation* condition (see Figures 18) were also included in this experiment. The ingroup condition was used to control for the fact that toddlers might just expect the target individual to hinder anyone whom they had witnessed hindering others, regardless of their group status. The outgroup-no-provocation condition was included to control for the fact that toddlers might expect the target individual to hinder any outgroup member.
Method

Participants

Participants were 48 healthy toddlers (23 male and 25 female) from English speaking families (27 months, 22 day to 31 months, 10 days, \( M = 29 \) months, 10 days). Another 7 toddlers were tested but excluded, two because they were either drowsy (1) or distracted (1), and five because they had a looking time that was over 2.5 standard deviations from the mean of their condition in the test trial. Sixteen toddlers were randomly assigned to the outgroup condition \( (M = 29 \) months, 15 days), the ingroup condition \( (M = 29 \) months, 6 days), and the outgroup-no-provocation \( (M = 29 \) months, 6 days) condition. In each condition, half the toddlers saw the help event, and half saw the hinder event.

Apparatus

The apparatus was similar to the one used in Experiments 1-5, but the parameters were slightly different. The apparatus consisted of a wooden display booth (127.5 cm high \( \times \) 101 cm wide \( \times \) 73.5 cm deep) mounted 76 cm above the floor of a brightly lit test room. The child sat on a parent’s lap and faced a large opening (44 cm \( \times \) 93.5 cm) in the front of the apparatus; between trials, a muslin-covered frame (61 cm \( \times \) 99.5 cm) was lowered to hide this opening. The back and side walls of the apparatus were painted white; the floor extended 11.5 cm behind the back wall and was covered with granite-patterned contact paper. Two wooden frames, each 192 cm high, 69 cm wide, and covered with white cloth, stood at an angle on either side of the apparatus; these frames served to isolate the child from the testing room and to hide the observers sitting behind them.

Three female native speakers of English were involved in the events. A1 wore a grey shirt and knelt at the right window; B1 wore a maroon shirt and knelt at the left window; the target individual wore a brown and beige shirt and sat at the back window.

Stimuli

At the beginning of the first familiarization trial, a transparent pink plastic cup (10.5 cm \( \times \) 8.5 cm in diameter at its largest points) was placed upside down on the apparatus floor 20 cm in front of the back
window and 7 cm from the left window. The cup had a transparent green plastic lid (0.5 cm X 9 cm in diameter) which was placed on the apparatus floor 37 cm in front of the back window and 7 cm from the left window. A grey granite-patterned placemat (0.5 cm X 10.5 cm X 10.5 cm) was placed on the apparatus floor 34.5 cm in front of the back window and 24 cm from the left window. Centered on the placemat was a wooden cube (2 cm X 2 cm X 2 cm) covered with solid green contact paper.

At the beginning of the second familiarization trial, a transparent orange plastic cup (10.5 cm X 8.5 cm in diameter at its largest points) was placed upside down on the apparatus floor 20 cm in front of the back window and 7 cm from the left window. The cup had a transparent yellow plastic lid (0.5 cm X 9 cm in diameter) placed on the apparatus floor 37 cm in front of the back window and 7 cm from the left window. A grey granite-patterned placemat (0.5 cm X 10.5 cm X 10.5 cm) was placed on the apparatus floor 34.5 cm in front of the back window and 68 cm from the left window. Centered on the placemat was a wooden cube (2 cm X 2 cm X 2 cm) covered with solid blue contact paper.

At the beginning of the test trial, a puzzle board (1 cm X 30 cm X 30 cm at its largest points) was placed on the apparatus floor 2 cm in front of the back window and 2 cm from the right window. In total, 5 puzzle pieces of different sizes and colors could fit into the puzzle board. A metal tray (1 cm X 25 cm X 20 cm) covered with blue contact paper was placed on the apparatus floor 34 cm in front of the back window and 1 cm from the right window. In the tray was a yellow square, an orange circle, a red rectangle and a blue triangle puzzle piece. A green oval puzzle piece was placed on the apparatus floor 20 cm in front of the back window and 7 cm from the left window, out of A1’s reach.

Trials

Familiarization trials

In the outgroup condition, the first familiarization trial only had an initial phase, which lasted 42 s. After a pause (2 s), A1 announced, “I’m a jaybo.” (2 s); after a 1-s pause, the target individual announced, “I’m a topid.” (2 s); then after a 1-s pause, B1 announced, “I’m a topid, too.” (2 s) The labels were repeated twice. The order of labeling was counterbalanced across participants (half of the children saw the labeling start from the A1, and half saw the labeling start from the B1). After the labeling
segment, B1 made a rattle by placing the block inside a cup (11 s) and shaking it (11 s) while A1 and the
target individual looked on.

The initial phase of the second trial thus lasted 54 s and was similar to the first familiarization trial except that the placemat lay across the apparatus floor near A1. Therefore, after the labeling, B1 attempted but failed to reach for the block (9 s). Then A1 took the block away and left the apparatus by closing the curtain behind her (6 s). B1 looked down and appeared unhappy (19 s).

In the ingroup condition, the two familiarization trials were similar to those in the outgroup condition except that at the beginning of each familiarization trial, A1 first announced, “I’m a jaybo.” (2 s); after a 1-s pause, the target individual announced, “I’m a jaybo, too.” (2 s); then after a 1-s pause, B1 announced, “I’m a topid.” (2 s)

In the outgroup-no-provocation condition, the two familiarization trials were similar to those in the outgroup condition, except each trial ended after A1, the target individual and B1 had each labeled themselves twice and thus lasted 20 s.

Test event

The test trial consisted of an initial and a final phase. The initial phase lasted for 29 s. B1 was absent (the left window was completely closed with an opaque white curtain), while A1 and the target individual were present. After a pause (2 s), A1 put all the puzzle pieces into the puzzle board except for one piece (12 s). She vainly attempted to reach for the last puzzle piece which lay across the apparatus out of reach (6 s). Then A1 left the apparatus by closing the curtain behind her (2 s). After A1 was gone, the target individual looked at the last puzzle piece (2 s). In the help event, the target individual put the last puzzle piece in A1’s tray with her right hand (4 s). She then withdrew her hand on the apparatus floor, looked at a neutral point, and paused (1 s). In the hinder event, the target individual threw the last puzzle piece out of the apparatus with her right hand (4 s). She then withdrew her hand on the apparatus floor, looked at a neutral point, and paused (1 s). During the final phase of the test trial, the toddlers watched this paused scene until the computer signaled the end of the trial. When this occurred, the supervisor lowered the curtain in front of the apparatus.
Procedure

The toddlers in the outgroup and the ingroup conditions were highly attentive and looked for 41.1/42 s on average in the first familiarization trial and for 52.6/54 s in the second familiarization trial. The toddlers in outgroup-no-provocation condition were also highly attentive and looked for 19.7/20 s on average in the first familiarization trial and for 19.7/20 s in the second familiarization trial.

Next, each toddler watched a single test trial in each condition, half the toddlers saw the help event, and half saw the hinder event. The toddlers in the three conditions were highly attentive during the initial phase of the test trial and looked for 28.6/29 s on average. The final phase ended when the toddler (1) looked away for 0.5 consecutive seconds after having looked for at least 10 cumulative seconds, or (2) looked for 60 cumulative seconds without looking away for 0.5 consecutive seconds.

The primary observer was absent from the test room during the familiarization trials and so did not know whether the toddler was assigned to the outgroup, the ingroup, or the outgroup-no-provocation condition. Interobserver agreement during the final phase of the test trial was calculated for 45/48 toddlers (only one observer was present for the other toddlers) and averaged 91% per toddler.

Preliminary analyses of toddlers’ looking times during the final phase of the test trial revealed no interaction of sex and order of labeling (whether the labeling started from A1 or B1) with condition (outgroup, ingroup, or outgroup-no-provocation) and event (help or hinder), $F$s < 1.69, $p$s > .05; the data was therefore collapsed across sex in the final analyses.

Results and Discussion

Toddlers’ looking times (see Figure 19) during the final phase of the test trial were compared by an ANOVA with condition (outgroup, ingroup, or outgroup no-provocation) and event (help or hinder) as between-subjects factors. The interaction between condition and event was significant, $F(2, 42) = 10.52, p < .001, \eta^2_p = 0.33$, but the main effects of condition and event were not, $F$s < 1.09, $p$s > .05. Planned comparisons revealed that in the outgroup condition, the children who saw the help event ($M = 27.0, SD = 9.1$) looked reliably longer than those who saw the hinder event ($M = 14.1, SD = 4.7$), $F(1, 42) = 9.13, p < .05, \text{Cohen’s } d = 1.78$; in the ingroup condition, the reversed pattern was found: the children who saw the
hinder event \((M = 24.9, SD = 10.0)\) looked reliably longer than those who saw the help event \((M = 10.4, SD = 2.7)\), \(F(1, 42) = 11.53, p < .05,\) Cohen’s \(d = 1.98;\) in the outgroup-no-provocation condition, the children who saw the hinder \((M = 18.1, SD = 14.4)\) and those who saw the help \((M = 14.2, SD = 4.2)\) event looked about equally, \(F(1, 42) < 1.\)

These results are consistent with those found in Experiment 4, and as such provide converging evidence for an early expectation of co-retaliation. Similar to the infants, the toddlers in the outgroup condition expected the target individual to hinder A1, who has previously hindered the target individual’s ingroup member, not because the toddlers expected the target individual to hinder anyone who has hindered, nor because the toddlers expected the target individual to hinder any outgroup member.

Rather, both infants and toddlers appear to attend to group membership and expect individuals (1) to co-retaliate against outgroup individuals who have acted negatively toward ingroup members, and also (2) to help ingroup members in need of assistance irrespective of whether they have acted negatively toward outgroup individuals. These expectations are not specific for a particular type of social group, but are probably general expectations about how individuals should act among social groups, including those that are arbitrarily marked by physical similarities or by labels.

These findings point to a consistent and continuous picture of early socio-moral understanding. Even before entering preschool, which marks the beginning of a typical period of intense peer interactions, infants and toddlers already have rich expectations about how individuals should interact within and across social groups. They expect individuals to protect group members through co-retaliation and to provide group members with assistance when needed. These expectations emerge in infancy and are maintained into toddlerhood.

**Experiment 7**

In the outgroup condition of Experiments 4 through 6, the children expected the target individual to hinder when the victim of the provocation was her ingroup member and to help when the perpetrator of
the provocation was her ingroup member. But what if the offender and the victim were both ingroup members of the target individual? Would the children expect the target individual to hinder or help?

In Experiment 7, we tested toddlers in an all-ingroup condition in which all three experimenters shared the same naming label (see Figures 20 & 21). The target individual was then faced with a dilemma: should she provide assistance to an ingroup member who had hindered another ingroup member? Or would she betray her ingroup loyalty by helping someone who had acted negatively toward the other ingroup member? We predicted that toddlers expect the target individual to hinder rather than help her ingroup member A1, who had hindered another ingroup member. As previously mentioned, the third-party punishment that can be enacted in this situation is critical for promoting prosocial behavior within groups, and for preventing the violation of social norms.

Prior research on third-party punishment suggests that people are willing to punish non-cooperators in their groups even when they themselves do not directly benefit from the punishment. For example, in a study by Gummerum and colleagues (2009), adults’ and children’s tendencies to punish non-cooperators were measured in a third-party punishment game, in which players could pay money from their own initial endowment to punish another anonymous unfair player. All participants were presented with an unfair scenario: the “proposer” allocated 18 of 20 Euros for him/herself, and left only 2 Euros for the “responder” in a dictator game. Participants began with an initial personal endowment of 10 Euros. Participants could pay any amount from 0 to 10 Euros, and this sum would be subtracted from their endowment. However, twice that amount would be subtracted from the payoff of the “proposer”. In the ingroup condition, the “proposer”, the “responder” and the participant all belonged to the same group (marked by their estimation in the dot estimation task). In the outgroup condition, the “proposer” and the “responder” belonged to the same group, but the participant belonged to a different group. The results showed that adults’ punishment for the unfair “proposer” was more than twice as high on average in the ingroup than in the outgroup condition. Similarly, sixth-graders tended to punish the unfair “proposer” more severely when they came from the ingroup than from the outgroup. Therefore, these results suggest that when a moral transgression occurs within their own groups (i.e., both the victim and the perpetrator
of the transgression are ingroup members), adults and elementary-school children are willing to enact third-party punishment even at their own cost; however, when both the victim and the perpetrator are outgroup members, they are less likely to scarify their personal interest to punish the perpetrator.

Because this third-party punishment game entails not only the difficulties of understanding and remembering complicated game rules and also the conflicts of self interest, it is probably hard to adopt this task to test younger children’s tendency to seek justice within their own group. However, younger children may also reveal their expectations about third-party punishment in simpler situations. In Experiment 7, toddlers observed, as bystanders, simple interactions among three members of a group, and their expectations about the interactions were measured by their looking times. These features of our task could reduce the information-processing load on toddlers, which might make them more likely to succeed at our task.

We predicted that the toddlers in Experiment 7 would show an expectation of third-party punishment within groups and thus would expect the target individual to hinder rather than help her ingroup member A1, who had hindered another ingroup member.

Method

Participants

Participants were 16 healthy toddlers from English speaking families, 8 male and 8 female (28 months, 14 day to 36 months, 11 days, $M = 30$ months, 10 days). Another toddler was tested but excluded because he had a looking time that was over 2.5 standard deviations from the mean of his condition in the test trial. Half the toddlers saw the help event, and half saw the hinder event.

Apparatus and Stimuli

The apparatus and the stimuli were identical to those used in Experiment 6.
Trials

Familiarization trials.

In the all-ingroup condition, the two familiarization trials were similar to those in the ingroup condition of Experiment 8, except that at the beginning of each familiarization trial, A1, on the right, first announced, “I’m a jaybo!” (2 s). After a 1-s pause, A2 (i.e., the target individual), in the middle, announced, “I’m a jaybo, too!” (2 s). After a 1-s pause, the A3, on the left, announced, “I’m a jaybo, too!” (2 s). The labels were repeated twice in each familiarization trial and the order of labeling was counterbalanced across participants.

Test event.

The test trial was identical to the one in Experiment 6.

Procedure

The procedure was similar to that in Experiment 6. The toddlers were highly attentive and looked for 41.4/42 s on average in the first familiarization trial and for 53.6/54 s in the second familiarization trial. The toddlers were also attentive during the initial phase of the test trial and looked for 29/29 s on average.

Interobserver agreement during the final phase of the test trial was calculated for 15/16 toddlers (only one observer was present for the other toddler) and averaged 95% per toddler.

Preliminary analyses of toddlers’ looking times during the final phase of the test trial revealed no interaction of sex and order of labeling (whether the labeling started from A1 or B1) with event (help or hinder), $F_{s} < 2.01, \ p_{s} > .05$; the data was therefore collapsed across sex in the final analyses.

Results and Discussion

Toddlers’ looking times (see Figure 19) during the final phase of the test trial were compared by a t-test with event (help or hinder) as the independent variable. In the all-ingroup condition the children who saw the help event ($M = 33.1, SD = 10.8$) looked reliably longer than those who saw the hinder event ($M = 16.2, SD = 4.7$), $t(15) = 4.42, \ p < .005$, Cohen’s $d = 2.03$. 

60
Next, we compared the looking times of the outgroup condition in Experiment 6 and those of the all-ingroup condition in Experiment 7. We found that the main effect of events (help or hinder) was significant, $F(1, 28) = 29.22, p < .05, \eta^2_p = 0.51$; the main effect of condition (outgroup or all-ingroup) was not significant, $F(1, 28) = 2.27, p > .05, \eta^2_p = 0.07$; the interaction between condition and event was not significant, $F(1, 28) < 1, p > .05$. Therefore, the toddlers in both the outgroup and the all-ingroup conditions showed the same looking pattern: they looked reliably longer at the help than at the hinder event.

These findings indicate that toddlers expect individuals to act negatively toward an ingroup member who has previously hindered another ingroup member. Also, the similar looking patterns of the outgroup condition and the all-ingroup condition suggest the expectations of co-retaliation is mostly triggered when someone acts negatively toward a group member, regardless of whether the perpetrator was another ingroup or an outgroup member. Our findings are consistent with the literature on third-party punishment. However, the ultimate functions of punishing ingroup versus punishing outgroup perpetrators still need further exploration. Punishing ingroup perpetrators may mainly help promote cooperation within groups, while punishing outgroup perpetrators may mainly help defend one’s groups. Nevertheless, the expectations of co-retaliation can conceptually encompass the expectations that an individual should retaliate against those who have acted negatively toward his/her ingroup members. Therefore, the targets of co-retaliation could be an outgroup or an ingroup perpetrator.

**Discussion of Project 2**

By the second year of life, children already attend to individuals’ group membership and hold different expectations for ingroup and outgroup members. Both infants and toddlers expect individuals (1) to co-retaliate against those who have acted negatively toward ingroup members, and (2) to help ingroup members in need of assistance irrespective of whether they have acted negatively toward outgroup individuals. Therefore, young children’s expectation of retaliation is modulated by considerations of ingroup loyalty, suggesting that they view retaliation as a retaliatory act not just for the individual’s self-interest, but also for the group’s interest.
Project 2 also provides evidence on infants’ expectations about the valence of individuals’ initial, as opposed to reciprocal, actions. Infants and toddlers expect individuals to act positively rather than negatively toward ingroup members, but they have no particular expectations about whether individuals should help or hinder an outgroup member without provocation. It was permissible to throw away the last puzzle piece (as in the outgroup-no-provocation condition of Experiments 4 and 6) or knock down a block of an outgroup member (as in the outgroup-no-provocation of Experiment 2). But it was not acceptable to do so toward an ingroup member. Therefore, infants and toddlers expect more positive initial actions within than across groups.

Project 2 suggested that infants and toddlers are already sensitive to social groups signaled by novel artificial group markers. In this project, we examined infants’ sensitivity to physical similarity (i.e., salient outfits) and also toddlers’ sensitivity to naming labels. Based merely on these transient and surface features of the individuals, infants form social groups and have different expectations for ingroup and outgroup individuals. Given that the social groups used in these experiments were novel and artificial in nature, these expectations are not likely to be specific to a particular type of social group, but are probably broad expectations about how individuals should act among social groups in general.

Furthermore, infants are not only aware of individuals’ group membership, but they also attribute differentiated mental states to each member of a group. Prior research on early psychological reasoning has already found that when infants watch an individual act in a scene, the psychological-reasoning system provides infants with some core concepts about the individual’s mental states. Infants can interpret the individual’s actions in terms of the following: motivational states, which specify the individual’s motivation in the scene (e.g. goals, dispositions); informational states, which specify what accurate information the individual can gather about the scene through perception, memory, or inference (e.g. knowledge, ignorance); and fictional states, which specify whether the individual’s representation of the scene is incompatible with that of the infant (e.g. false belief, pretense). The cognitive subsystem that enables infants to attribute motivational and informational states is already operational in the first months of life and is well in place by the end of the first year (e.g. Csibra, 2008; Johnson, et al., 2007; Liszkowski,
et al., 2006; Luo & Baillargeon, 2005, 2007; Luo & Johnson, 2009; Song & Baillargeon, 2007; Tomasello & Haberl, 2003; Woodward, 1998). The findings from Project 2 extend our understanding of early social cognition. When reasoning about interactions within or across social groups, infants do not solely consider group members as tokens of the group; infants can interpret and predict group members’ actions based on the mental states of the group member, as well as on a set of socio-moral norms.

The findings from Project 2 also provide evidence for the emergence and the continuity of early socio-moral expectations. Even before experiencing rich peer interactions in the preschool years, infants and toddlers already have rudimentary expectations about how individuals should interact with ingroup and outgroup members. Infants can interpret group interactions not only in dyadic, but also in triadic form, and they are able to consider immediate, as well as prior interactions. Infants understand that individuals’ behavior is affected not only by actions directed at themselves, but also by actions directed at their ingroup members. These early socio-moral expectations probably emerge early in life, and will be consolidated or modified by later experience. Future investigation is needed to examine the relations between early expectations and children’s reactions when they begin to engage in increasingly sophisticated social interactions in the preschool and early school years.
CHAPTER 4: PROJECT 3 ON PRIVILEGE

Projects 1 and 2 focused on how ingroup loyalty modulates early expectations about retaliation (i.e., negative reciprocity). In Project 3, we examined whether the expectations about reciprocation (i.e., positive reciprocity) might also be modulated by considerations of ingroup loyalty. We also explored whether children expect reciprocation in both ingroup and outgroup contexts, or if they expect it only in the ingroup context.

Reciprocation (e.g. “I’ll scratch your back if you scratch mine”) is a critical feature of social interactions within groups. A social group can be thought of as a “container” of direct and indirect reciprocity. People tend to favor ingroup over outgroup members and they expect similar favorable treatment in return (e.g. Yamagishi & Kiyonair, 2000; Yamagishi et al., 1999). Studies on chimpanzees’ grooming behavior showed that chimpanzees groom with one another regularly within their groups. However, this grooming behavior rarely occurs between outgroup members. Chimpanzees use reciprocal grooming to maintain cooperative relationships, which ultimately influences their dominance rank and access to mates (e.g., Watts, 2000).

In human societies, reciprocity is commonly expected within groups, but to a lesser extent across groups. For example, in a sequential prisoner’s dilemma game, groups were arbitrarily formed based on trivial criteria, such as participants’ preferences for one painter over another. Adults allocated more resources to an ingroup member than to an outgroup member. They also expected ingroup members to show similar favorable allocations in return (e.g., Yamagishi et al., 1999).

Children themselves tend to reciprocate actions of an ingroup member to a greater degree than an outgroup member. The study by Gummerum and colleagues (2009) also investigated the reciprocation tendencies of adults and elementary-school children in a sequential prisoner’s game. In this game, two anonymous players have an initial monetary endowment, which they can sequentially exchange between each other. The money they give to the other player from their endowment is doubled and added to the other’s endowment. In this study, all participants played as the second player. For example, both players initially had an endowment of 5 Euros each. Participants, as second players, were given all 5 Euros from
the first player, doubled. Thus, they possessed 15 Euros altogether before they decided how much to give back to the first player from their 15 Euros. Whatever amount participants gave to the first player would be doubled and added to the first player’s account. The players’ group membership was marked by their estimations in the dot estimation task. The results showed that adults reciprocated the first player’s cooperative act equally with ingroup and outgroup players. However, children from sixth grade reciprocated more with ingroup than with outgroup players. Naturalistic observations in preschool settings have also shown that 3- and 4-year-olds reciprocate object offering more often with friends than with non-friends during free play time (Fujisawa, Kutsukake, & Hasegawa, 2010).

Project 3 examined early expectations about *ingroup privilege* and asked whether young children expect group members to have the privilege of being reciprocated more positively. We hypothesized that (1) toddlers would expect individuals to reciprocate *both* their ingroup and outgroup members; however, (3) given their considerations of ingroup loyalty, they would expect individuals to reciprocate *more positively* in response to a positive overture by an ingroup than by an outgroup member.

We tested whether toddlers would expect an individual to reciprocate a positive overture from an ingroup member (Experiment 8) or from an outgroup member (Experiment 9). By comparing the results of Experiments 8 and 9, we examined whether toddlers expect a positive act to be reciprocated with a *greater magnitude* within compared to across groups.

Finally, in Experiment 10, we tested whether toddlers would expect an individual to reciprocate a positive act when group membership was *unmarked*. This experiment was intended to help understand whether young children’s expectations of ingroup privilege are due to ingroup favoritism or outgroup derogation.

**Experiment 8**

Experiment 8 tested whether 3-year-old toddlers would expect an individual to reciprocate a positive overture from an ingroup member; and if so, then what magnitude would be viewed as appropriate. Membership was marked by giving the individuals novel naming labels.
Toddlers watched two familiarization trials and one test trial (see Figure 22). At the start of each familiarization trial, three individuals sat at an apparatus: A1, on the right; B1, on the left; and the target individual (who acted towards A1 in response to her positive overture in the test trial) sitting in the middle. The target individual and A1 belonged to the same group while B1 to another group. Group membership was identified by naming labels: in an initial trial, A1 announced, “I am a boga!” The target individual announced, “I am a boga, too!” Then B1 announced, “I am a mayloo!” The labels were repeated twice in each familiarization trial, with order counterbalanced across participants.

In test, B1 was absent, and A1 asked the target individual to play with her, offering her one of two identical toy penguins. In the refuse event, the target individual crossed her arms and looked away from A1, who kept tilting her toy left and right until the trial ended. In the accept event, the target individual picked up the toy offered by A1, looked at it, then put it down and paused while looking at it until the trial ended. In the coordinate event, the target individual picked up the toy offered by A1, and then they both tilted their toys left and right in unison until the trial ended.

We speculated that infants would expect the target individual to act positively towards A1 in response to her positive overture, and thus would view both the accept and the coordinate events as acceptable, but not the refuse event.

Method

Participants

Our design called for 24 participants. Thus far, 22 healthy 3-year-old toddlers (10 male and 12 female from English speaking families have been tested (33 months, 0 day to 37 months, 3 days, $M = 34$ months, 26 days). Another 8 toddlers were tested but excluded, one because of parental interference, and seven because they looked for the maximum allowed in the test trial. One third of the toddlers were randomly assigned to the refuse ($M = 35$ months, 1 days), the accept ($M = 34$ months, 12 days), or the coordinate event ($M = 35$ months, 3 days).
Apparatus and Stimuli

The apparatus was the same as in Experiments 6 and 7.

Three female native speakers of English were involved in the events. A1 wore a blue shirt and knelt at the right window; B1 wore a green shirt and knelt at the left window; the target individual wore a red shirt and sat at the back window.

During the test trial, A1 brought in two identical plastic penguins toys. Each penguin toy (6 cm X 6.5 cm X 6.5 cm at its largest points) had a yellow body, black eyes, a white beak, pink wings, and a white stomach decorated with a yellow fish.

Trials

Familiarization trials

The first familiarization trial only had an initial phase, which lasted 20 s. The trial began with a brief pause (2 s), followed by A1 announcing, “I’m a boga.” (2 s); after a 1-s pause, the target individual announced, “I’m a boga, too.” (2 s); then after a 1-s pause, B1 announced, “I’m a mayloo.” (2 s) The labels were repeated twice. The order of labeling was counterbalanced across participants (half of the children saw the labeling start with A1, and half saw the labeling start with B1). Next, the computer signaled the end of the event, and the supervisor lowered the curtain in front of the apparatus.

The second familiarization trial was identical to the first familiarization trial.

Test trial

The test trial consisted of an initial and a final phase. The initial phase lasted about 16 s. B1 was absent and the left window was completely closed with an opaque white curtain. After a pause (2 s), A1 held up both toys and oriented them toward the target individual (2 s). A1 tilted both toys left and right in unison (4 s). Then A1 placed both toys down on the apparatus floor, one in front of herself and the other in front of the target individual (1 s). After a pause (1 s), A1 grasped the toy in front of herself, held it up, and oriented it toward the target individual (2 s). A1 then started to tilt her toy left and right (1 s).
In the *refuse* event, the target individual crossed her arms, looked away from A1 and paused (3 s). During the final phase, the target individual paused with her arms crossed and her head turned away, while A1 continued tilting her toy left and right until the event ended.

In the *accept* event, the target individual picked up the toy offered by A1 and held it in front of herself with both hands and looked at it (2 s). She then put it down on the apparatus floor, withdrew both hands on her lap, looked down at the toy, and paused, while A1 continued tilting her toy left and right in unison until the trial ended (1 s).

In the *coordinate* event, the target individual picked up the toy offered by A1 with her right hand (1 s) and tilted the toy to left and right (2 s). During the final phase, both the target individual and A1 kept tilting their toys left and right in unison until the trial ended.

**Procedure**

The toddlers were highly attentive during the familiarization trials and looked for 20/20 s on average.

Next, each toddler watched a single test trial; one third of the toddlers saw the refuse event, one third saw the accept event, and one third saw the coordinate event. The toddlers were attentive during the initial phase and looked for 15.9/16 s on average. The final phase ended when the toddler (1) looked away for 0.5 consecutive seconds after having looked for at least 10 cumulative seconds, or (2) looked for 60 cumulative seconds without looking away for 0.5 consecutive seconds.

The primary observer was absent from the test room during the familiarization trials and therefore did not know whether the infant was assigned to the ingroup (Experiment 8) or the outgroup condition (Experiment 9). Interobserver agreement during the final phase of the test trial was calculated for 20/22 infants (only one observer was present for the other infants) and averaged 93% per toddler.

Preliminary analyses of toddlers’ looking times during the final phase of the test trial revealed no interaction of sex and order of labeling (whether the labeling started from A1 or B1) with event (refuse, accept or coordinate), $F_s < 3.04$, $p_s > .10$; the data was therefore collapsed across sex and order of labeling in the final analyses.
Results and Discussion

Toddlers’ looking times (see Figure 23) during the final phase of the test trial were compared by an ANOVA with event (refuse, accept, coordinate) as a between-subjects factor. The effect of event was marginally significant, $F(2, 19) = 3.30, p = .059, \eta^2_p = 0.19$. Planned comparisons indicated that the toddlers who saw the refuse event ($M = 25.3, SD = 3.6$) looked reliably longer than those who saw either the accept ($M = 18.1, SD = 5.6$) or the coordinate event ($M = 18.6, SD = 7.5$), $F_s(1, 19) > 4.86, ps < .05$; the looking times at the accept and the coordinate events were approximately the same, $F(1, 19) < 1$.

The toddlers thus expected the target individual, following her ingroup member A1’s friendly overture, to reciprocate either by simply accepting the toy or by playing with the toy in coordinated rhythm with A1. Also, the toddlers considered it unacceptable to refuse A1’s friendly offer by crossing her arms and turning away. These findings suggest that toddlers expect individuals to reciprocate a positive act by an ingroup member, and that they consider refusal as negative and as an inappropriate response to an ingroup member’s friendly overture.

Experiment 9

Experiment 9 tested whether 3-year-old toddlers would expect the target individual to reciprocate a positive overture made by an outgroup member. In addition, we examined whether the toddlers would expect the target individual to respond to a positive action by an outgroup member less positively compared to an ingroup member.

The findings from Projects 1 and 2, as well as prior studies (e.g., Over & Carpenter, 2009) suggest that children expect the initial (as opposed to reciprocal) action toward an ingroup member to be positive rather than negative, but that children have no particular expectation about whether the initial action toward an outgroup member will be positive or negative. In other words, the baseline for the valence of interactions in the ingroup versus outgroup condition may be more positive within than those across groups. Therefore, we speculated that the toddlers would expect different magnitudes of reciprocation from an ingroup versus an outgroup member.
The toddlers watched trials that were identical to those used in Experiment 8, except that the target individual and A1 belonged to different groups (see Figure 24). Therefore, in the beginning of each familiarization trial, A1 announced, “I am a boga!” The target individual announced, “I am a mayloo!” Then B1 announced, “I am a mayloo, too!”

We speculated that toddlers would expect the target individual (1) to reciprocate an outgroup member’s friendly overture; and (2) to respond less enthusiastically to the outgroup member’s friendly overture.

Method

Participants

Participants were 24 healthy 3-year-old toddlers from English speaking families, 12 male and 12 female (30 months, 12 day to 37 months, 22 days, $M = 34$ months, 18 days). Another 2 toddlers were tested but excluded, one because she was upset and the other one because he looked for the maximum allowed in the test trial. Eight toddlers were randomly assigned to the refuse ($M = 33$ months, 29 days), the accept ($M = 34$ months, 16 days), and the coordinate event ($M = 35$ months, 9 days).

Apparatus and Stimuli

The apparatus and stimuli were identical with those used in Experiment 8.

Trials

Familiarization trials

The first familiarization trial only had an initial phase, which lasted about 20 s. The trial began with a brief pause (2 s), followed by A1 announcing, “I’m a boga.” (2 s); after a 1-s pause, the target individual announced, “I’m a mayloo.” (2 s); then after a 1-s pause, B1 announced, “I’m a mayloo, too.” (2 s) The labels were repeated twice. The order of labeling was counterbalanced across participants; half of the children saw the labeling start from A1 while the other half saw the labeling start from B1. Next, the computer signaled the end of the event, and the supervisor lowered the curtain in front of the apparatus. The second familiarization event was identical to the first familiarization.
Test trial

The test trial was identical to that in Experiment 8 (each toddler watched a refuse, an accept, or a coordinate event).

Procedure

The procedure was identical to the one used in Experiment 8. The toddlers were highly attentive and looked for 19.4/20 s on average during the familiarization trials and for 16/16 s during the initial phase of the test trial. Interobserver agreement during the final phase of the test trial was calculated for 22/24 toddlers (only one observer was present for the other infants) and averaged 94% per toddler.

Preliminary analyses of toddlers’ looking times during the final phase of the test trial revealed no interaction of sex and order of labeling (whether the labeling started from A1 or B1) with event (refuse, accept or coordinate), $F$s < 1.02, $p$s > .39; the data was therefore collapsed across sex and order of labeling in the final analyses.

Results and Discussion

Toddlers’ looking times (see Figure 23) during the final phase of the test trial were compared by an ANOVA with event (refuse, accept or coordinate) as a between-subjects factor. The effect of event was significant, $F (2, 21) = 8.93, p < .005, \eta^2_p = 0.46$. Planned comparisons indicated that the toddlers who saw the coordinate event ($M = 28.3, SD = 8.5$) looked reliably longer than those who saw the accept ($M = 16.3, SD = 5.8$) or the refuse event ($M = 16.2, SD = 4.8$), $F$s (1, 21) > 13.39, $p$s < .005; the looking times at the accept and the refuse event were approximately the same, $F (1, 21) < 1$.

The toddlers’ looking times in Experiments 8 (ingroup condition) and 9 (outgroup condition) were analyzed by an ANOVA, with event and condition as two between-subject factors. The interaction between event and condition was significant, $F (2, 40) = 8.84, p < .005, \eta^2_p = 0.31$. The main effects of condition and event were not significant, $F$s <1. Planned interaction comparisons indicated that the toddlers who saw the refuse event in the ingroup condition looked reliably longer than those who saw the same event in the outgroup condition, $F (1, 40) = 7.94, p < .05, \eta^2_p = 0.17$; the toddlers who saw the coordinate event in the outgroup condition looked reliably longer than those who saw the same event in
the ingroup condition, $F(1, 40) = 9.67, p < .05, \eta^2_p = 0.19$; the looking times at the accept event were approximately the same in the ingroup and the outgroup conditions, $F(1, 40) < 1$.

The finding with the accept event was similar to that of the ingroup condition in Experiment 8. The toddlers expected the target individual to simply accept the toy offered by an outgroup member. Thus, toddlers viewed neutral acceptance of a friendly invitation as appropriate in both ingroup and outgroup interactions.

The results with the coordinate event suggested that the toddlers considered it unexpected for the target individual to engage in coordinate play with her outgroup member, A1. One possible explanation is that infants viewed playing in coordinated rhythm as a signature of within-group interactions. Prior evidence has shown that infants already expect members of a group to cohere and move together (e.g. Premack & Premack, 1997; Sloane et al., under review; Sugarman, 1983). Therefore, when the toddlers in Experiment 9 saw the target individual and A1 tilt the toys left and right in unison, the toddlers inferred that the two belonged to the same group. This was inconsistent with the toddlers’ initial categorization suggested by the novel labels. This inconsistency led the toddlers to reconsider the individuals’ group membership and hence to look longer at the coordinate event. Another possibility, though not contradictory to the first one, is that the toddlers viewed the coordinate event as too positive and beyond the range of reciprocatory action expected toward an outgroup member. In Project 1, we found that the infants had expectations about the range of retaliatory action in response to a provocation, and they detected a violation if the retaliation was outside the expected range, either too severe or too mild. Similarly, toddlers might have found the coordinate event beyond the upper limit of the range of reciprocatory action expected toward an outgroup member.

The results with the refuse event suggested that the toddlers considered it permissible for the target individual to refuse to play with an outgroup member. Would this not be a violation of reciprocity? As found in Projects 2 and 3, mild aggression toward outgroup members is tolerated without provocation, as though it is seen as part of the acceptable “neutral” range for interactions among individuals from
different groups. Therefore, when an individual receives a positive overture from an outgroup member, a refusal may be viewed as neutral rather than as negative for an outgroup member.

Overall, the results of this experiment suggest that the range of expected actions in response to a friendly overture is “shifted” towards the positive, compared to neutral situations. Because of the different baseline for the valence of interactions in the ingroup versus outgroup condition, reciprocatory action is expected to be more positive toward ingroup than toward outgroup members. The expected responses to an ingroup member’s friendly overture range from mildly positive actions (e.g., accepting the offered toy) to moderately positive actions (e.g., playing in coordinated rhythm), whereas the expected responses to a outgroup member’s positive overture range from mildly negative actions (e.g., refusing to play together) to mildly positive action (e.g., accepting the offered toy).

It is important to continue investigating the limit of negative action in response to a positive overture in the outgroup condition in the future. Our speculation is that toddlers might view throwing away the toy offered by the outgroup member as unacceptable, but view it as permissible without the friendly overture. This would also be supported by the results we found in the outgroup-no-provocation conditions of Project 3.

Therefore, toddlers expect individuals to follow the norm of reciprocation in both the ingroup and the outgroup conditions, in other words, toddlers expect individuals to act positively or at least moderately in response to positive act. But they expect individuals to reciprocate with a greater magnitude with ingroup members than with outgroup members.

Experiment 10

Experiment 10 examined whether 3-year-old toddlers would expect individuals to reciprocate a positive overture when group membership was unmarked. More specifically, we asked whether toddlers would consider it acceptable for the target individual to refuse playing with another individual whose group membership was unmarked, versus playing with that individual in a coordinated rhythm.

The toddlers watched a single test trial identical to the one in Experiments 8 and 9, but they did not know the labels of the three individuals (see Figure 25). In this test trial, the left experiment was
absent, and the right experimenter (RE) invites the target individual to play with her, offering her one of two identical toy penguins. In the refuse event, the target individual crossed her arms and looked away from RE, who kept tilting her toy left and right until the trial ended. In the coordinate event, the target individual picked up the toy offered by RE, and then they both tilted their toys left and right in unison until the trial ended. Given that the accept event was considered as acceptable in both the ingroup and outgroup conditions, we did not run the accept event in Experiment 10.

Prior research with adults indicates that ingroup bias is not a direct result of outgroup derogation, which reduces the level of cooperation toward outgroup individuals compared to neutral individuals whose group membership is unmarked; instead, ingroup bias results from ingroup favoritism, which increases the level of cooperation toward ingroup individuals compared to neutral individuals (Brewer, 1979, 1999; Koopmans & Rebers, 2009, Yamagishi & Mifune, 2009). For example, in a Prisoners’ Dilemma Game, Japanese adult participants contributed more to an ingroup partner than to an outgroup or a neutral partner whose group membership was unknown. In other words, the participants did not treat the outgroup partner and the neutral partner differently, but they did treat their ingroup partner more positively (e.g., Koopmans & Rebers, 2009). Moreover, when group membership is private knowledge (i.e., unknown for other group members), participants cannot expect in-group favoring behavior from the in-group member; therefore, bias toward ingroup members is no longer displayed (e.g., Yamagishi et al., 2005, 2008). Therefore, it is necessary to know the two players’ group membership in order for the in-group bias to emerge.

We speculated that the toddlers would treat the unmarked condition the same as the outgroup condition. Therefore, toddlers in the unmarked condition would look longer at the coordinate than at the refuse event, just as in the outgroup condition of Experiment 9.

Method

Participants

Our designed called for 16 participants. So far, we have texted 13 healthy 3-year-old toddlers (7 male and 6 female) from English speaking families (32 months, 2 day to 35 months, 13 days, $M = 34$
Another 4 toddlers were tested but excluded, one because he had a looking time that was over 2.5 standard deviations from the mean of the condition in the test trial, and three because they looked for the maximum allowed in the test trial. Half of the toddlers saw the refuse event ($M = 34$ months, 3 days) and half saw the coordinate event ($M = 34$ months, 18 days).

**Apparatus and Stimuli**

The apparatus and stimuli were identical to those used in Experiments 8 and 9.

**Trials**

Experiment 10, unlike Experiments 8 and 9, included no familiarization trial because no labels were announced to indicate the group membership.

**Test trial**

Each toddler received only a single test trial in which they saw either a refuse or a coordinate event. These were identical to those used in Experiments 8 and 9.

**Procedure**

The procedure was identical to the one used in Experiments 8 and 9. The toddlers were highly attentive during the initial phase and looked for 15.9/16 s on average in the test trial. Interobserver agreement during the final phases of the test trials was calculated for 12/13 infants (only one observer was present for the other infant) and averaged 95% per trial per infant.

Preliminary analyses of toddlers’ looking times during the final phase of the test trial revealed no interaction of sex with event (refuse or coordinate), $F(1, 9) < 1$. The data was therefore collapsed across sex in the final analyses.

**Results and Discussion**

Toddlers’ looking times (see Figure 26) during the final phase of the test trial were compared by a $t$-test with event (refuse or coordinate) as an independent variable. The effect of event was marginally significant, $t (11) = 2.05, p = .06$, suggesting that the toddlers who saw the coordinate event ($M = 25.5, SD = 8.2$) looked longer than those who saw the refuse event ($M = 19.3, SD = 2.6$), Cohen’s $d = 1.02$. 
The toddlers’ looking times at the coordinate event and the refuse event in Experiments 8 (ingroup condition), 9 (outgroup condition), and 10 (unmarked condition) were analyzed by an ANOVA, with event and condition as between-subjects factors. The interaction between event and condition was significant, $F(2, 38) = 9.14, p < .005, \eta^2_p = 0.35$. Planned interaction comparison indicated that the looking patterns of the ingroup and outgroup conditions were significantly different, $F(1, 38) = 17.91, p < .05, \eta^2_p = 0.32$; the looking patterns of the ingroup and unmarked conditions were also significantly different, $F(1, 38) = 15.45, p < .05, \eta^2_p = 0.29$; however, the looking patterns of the outgroup and unmarked conditions were not significantly different, $F(1, 38) = 1.59, p > .05, \eta^2_p = 0.04$.

These findings show that the toddlers treated the outgroup and unmarked conditions similarly. This suggests that the toddlers expected the individuals to respond more positively and enthusiastically in the ingroup condition than in either the outgroup or the unmarked condition.

**Discussion of Project 3**

Consistent with the results from Project 2, the toddlers in Project 3 attended to individuals’ group membership marked by novel naming labels, and they held different expectations for ingroup and outgroup members. The results of Project 3 suggest that early expectations about reciprocation, similar to those about retaliation, are modulated by considerations of ingroup loyalty. Toddlers, by 3 years, already expect ingroup members to have the privilege of being reciprocated more positively than outgroup members.

Is this privilege rooted in out-group hatred or ingroup love? The findings of Project 3 suggest that toddlers treat the outgroup condition as equivalent to the neutral, unmarked condition. The difference in the expected magnitude of reciprocation between the ingroup and outgroup conditions is due to an increase in magnitude of reciprocation for ingroup members, but not due to a decrease in magnitude of reciprocation for outgroup members. Therefore, the expected ingroup privilege is primary caused by ingroup love rather than outgroup hatred. These findings are consistent with the claim in the adult literature that people have no intrinsic “outgroup hatred,” but only “ingroup love” (e.g. Brewer, 1979).
The early expectation of ingroup privilege is probably a byproduct of social exchange and sharing that takes place within the boundaries of a group. On the one hand, there are potentially more interactions with ingroup members in daily lives. Immediate reciprocation is important for within-group cooperation to arise, which promote reciprocation in the future (e.g. Yamagishi & Kiyonair, 2000; Yamagishi et al., 1999). The positivity within a group is reinforced by the frequent and recurring interactions with group members. On the other hand, interactions with outgroup members is relatively infrequent, and hence individuals have less incentives to reinforce future positive actions by outgroup members. In addition, coordination with an outgroup member may violate one’s sense of ingroup loyalty which can be costly (e.g., Brewer, 1999). This idea was also supported by the results of the coordinate events in Experiments 9 and 10. Therefore, the expected ingroup privilege is a natural outcome of our constrained social circles. In modern societies with increases in travel and access to information as well as political and economical interdependence, we have more contacts with outgroup members than our ancestors, and thus our tendency to grant privilege to group members may be attenuated.

The results from Projects 2 and 3 together suggest that 3-year-old children consider naming labels sufficient to establish group membership, even when the labels are just nonsense nouns. This confirms the robust effect of labeling on social categorization (e.g., Bigler, 1995; Bigler, Jones, & Loblinier, 1997; Bigler, Brown, & Markell, 2001). These results also suggest that social categorization is malleable at least in early years. In our experiments, the experimenters varied naturally in the color of their skins, hair, eyes, and so on. Three-year-olds categorizes the experimenters based on their announced labels. This indicates that children may be able to re-categorize people of different races, genders, and ages based on new labels. This flexibility may be evolutionarily adaptive for children to infer, from environmental cues, which bases of categorization are important within a given context (Bigler & Liben, 2007).
CHAPTER 5: GENERAL DISCUSSION

The norm of reciprocity, including both retaliation and reciprocation, is widely considered to be one of the fundamental norms guiding human social interactions (e.g., Fehr & Gächter, 2000; Fry, 2006; Gouldner, 1960; Jackendoff, 2007; Premack, 1990). The results of the present projects show that infants and toddlers have general expectations about reciprocity both within and across groups. Moreover, their expectations of retaliation and reciprocation are modulated by considerations of ingroup loyalty. Specifically, infants and toddlers have the following expectations for social interactions: (1) more severe retaliation against outgroup than against ingroup members (expectation of escalation); (2) retaliation against outgroup members who have acted negatively toward an ingroup member other than oneself (expectation of co-retaliation); and (3) more positive reciprocation with ingroup than with outgroup members (expectation of privilege).

As shown in the proposed model (see Figure 27), the range of expected actions shifts based on two major factors: (1) whether the two individuals belong to the same or different groups (ingroup vs. outgroup condition); (2) whether there has been a prior initial action, and if so, whether the reciprocal action occurs in response to a negative initial action (i.e., retaliatory action) or to a positive initial action (i.e., reciprocatory action). The scale of the diagram is ordinal rather than ratio (e.g., the distance between mildly and moderately negative may not be exactly the same as between moderately and extremely negative).

The results of the present projects suggest that the baseline for an initial action between ingroup members is more positive than between outgroup members. Infants expect individuals to act positively toward ingroup members while they have no particular expectations about whether initial actions toward outgroup members should be (mildly) positive or negative. In reality, the range of acceptable actions toward ingroup versus outgroup members probably varies depending on the nature of the specific group membership. The initial actions toward family members are expected to be more positive whereas the initial actions toward war enemies are expected to be more negative. Here, the diagram depicts a general
range of acceptable initial actions among members of horizontal and moderate social groups, which involve neither group hierarchy nor personal intimacy.

Furthermore, the range of acceptable retaliatory actions are expected to be more negative than the range of acceptable initial actions, with a “shift” toward the negative end of the spectrum in both the ingroup and outgroup conditions. Of course, the degree of the “shift” toward the negative end also depends on the severity of the provocation (e.g., from stealing a cracker to threatening one’s life). Importantly, because of the different baselines, retaliatory actions are relatively more negative in outgroup than in ingroup conditions. Mild negative actions toward outgroup members are not inhibited in the first placed and then moderate retaliation are expected in response to those negative actions, which eventually caused spiral escalation of conflicts among groups.

Correspondingly, the range of acceptable reciprocatory actions may be expected to be more positive than initial actions in both the ingroup and the outgroup conditions. Reciprocation may result in a “shift” toward the positive end of the spectrum for these interactions, but because of the different baselines, reciprocatory actions are relatively more positive in the ingroup than in the outgroup condition. The distance of the “shift” toward the positive end depends on the magnitude of the initial positive act (e.g., from offering a toy to saving one’s life). In Project 3, A1’s friendly overture (i.e., offering a toy to play) was mildly positive, and therefore the “shift” was quite subtle. Future investigation on the limits of responses to a positive act is needed. For example, evidence that toddlers view throwing away the toy offered by the outgroup member as unacceptable but view it as permissible without the friendly overture would provide evidence for the “shift” of reciprocatory actions.

In addition, we found that the expectations of escalation and co-retaliation were already present in the second year of life and the expectation of privilege was present in by at least the third year of life. We have not yet tested whether infants in the second year of life also expect that ingroup members should have the privilege of being reciprocated more positively. However, we believe that this expectation of ingroup privilege is also likely to emerge in infancy, because infants already show consistent favoritism to
ingroup over outgroup members (e.g., Kinzler, Dupoux, & Spelke, 2007; Mahajan & Wynn, 2010; Olson & Spelke, 2008; Shutts, Banaji, & Spelke, 2009; Werneken & Tomasello, 2006, 2007).

The presence of these socio-moral expectations is important, as they help young children interpret the social interactions they observe on a daily basis, anticipate the effects of their own actions toward others, and perhaps also help them to distinguish between individuals who are more or less likely to reciprocate positive actions or retaliate after negative interactions. As children grow up, they need to weather the storm of social situations, such as receiving a friendly invitation, being mistreated by a bully, or being called on for help. Early socio-moral expectations become part of the resources that guide children’s reasoning and behavior in these local social situations.

The results of present projects also provide insights that help formulate answers to the following questions:

1. What are early expectations about initiating behavior?

Similar to adults, infants and toddlers generally act more positively toward ingroup than toward outgroup individuals even without any prior interactions with those individuals (e.g., Kinzler, Dupoux, & Spelke, 2007; Mahajan & Wynn, 2010; Olson & Spelke, 2008; Shutts, Banaji, & Spelke, 2009; Werneken & Tomasello, 2006, 2007). This group bias is more likely to be driven by ingroup love than by outgroup hatred. Thus, outgroup hatred is less likely to be an intrinsic part of human nature than its counterpart, ingroup love (e.g., Brewer, 1999; Yamagishi & Mifune, 2009). The results of this series of experiments suggest that infants already have different expectations about actions initiated toward ingroup and outgroup members. These early expectations about people’s initial actions are important for establishing a baseline for understanding infants’ and toddlers’ expectations about intergroup interactions, because these group-specific set-points lead to differences in magnitude of reciprocation and retaliation as discussed above.

As an expansion of the idea that intergroup bias results from ingroup love rather than outgroup hatred, Yamagishi and colleagues (1999; 2000) suggest that a social group is a “container” of generalized reciprocity. People act positively within their own group and also expect other group members to treat
them positively. For example, in the study by Karp and colleagues (1993), Japanese participants were asked to divide 500 yen between an ingroup and an outgroup member (the group member was marked by overestimating or underestimating the number of dots in a dot estimation task). In addition, they were told that each participant would receive an allocation from both an ingroup and an outgroup, and their final payment would be the average of the two allocations. Results revealed that participants favored ingroup over outgroup members only when the participant was also the target of other participants’ allocation decisions. However, when participant’s own pay was fixed, and was not dependent on other participants’ allocation decisions, ingroup favoritism did not emerge. Ingroup favoritism also disappeared when the participants did not expect to receive a favorable allocation from other in-group members, or when their group membership was private (as opposed to common) knowledge (e.g., Yamagishi et al., 2005, 2008).

Infants’ ambivalent expectations about initial actions toward outgroup members, as found in Project 3, may be adaptive to the diverse intergroup relations in the world. In some pre-industrial cultures, hostile behavior toward outgroup members is praised and morally sanctioned; but in other cultures, which highly embrace egalitarianism, people tend to disapprove of outgroup hostility (e.g., Haidt & Joseph, 2007). Children have to observe and learn the best strategies to use when interacting with outgroup members. This flexibility can help children adapt to whatever culture they were born into.

2. How does ingroup loyalty manifest in reciprocal interactions?

The present projects provide evidence that infants and toddlers are already sensitive to a broad range of social group markers, from common social factors of race, gender, age, or accent, to arbitrary group markers based on clothing or novel names. In our experiments, infants and toddlers immediately formed different expectations for ingroup and outgroup individuals, even when they themselves were not affiliated with any of those groups, nor did they have prior exposure to those groups.

Infants’ considerations of ingroup loyalty are manifested in reciprocal interactions in several ways. Infants expect individuals to provide assistance to group members, to grant privilege to group members when reciprocating positive acts (e.g., engage in coordinated actions), to protect group members through co-retaliation, and to maintain harmony within a group by refraining from unprovoked aggression.
and escalated retaliation. Therefore, more positivity is expected for the initial actions toward ingroup than toward outgroup members; also, because of the different baselines, more positive reciprocation and less severe retaliation are expected within than across groups.

These expectations may encompass crucial facets of infants’ rudimentary understanding of ingroup loyalty. They may be some basic social rules that help infants reason about interactions within as compared to across groups. As pointed out by Hirschfeld (2001), understanding and acting according to the affordance of social groups is very important for children. Children have to know “who’s who” in a social environment, how to predict what other people are likely to do, and how to prepare interactions with both ingroup and outgroup members.

3. Is reciprocity the foundation of other socio-moral norms?

Reciprocity, the reward of kind acts and the punishment of unkind acts, is often described as one of the basic norms that guide social interactions (e.g., Fehr & Gachter, 2000; Stripada & Stich, 2007). Reciprocity constitutes a key aspect of morality in societies from around the world (e.g., Brown, 1991; Fry, 2006). In the Descent of Man, originally published in 1871, Charles Darwin proposes that reciprocity is the “foundation of morality.”

The present projects explored how the norm of reciprocity is modulated by the norm of ingroup loyalty, and found that infants and toddlers already have rich expectations about retaliatory and reciprocatory behaviors in intergroup contexts. These findings not only provide evidence for the breadth of early socio-moral expectations, they also raise interesting questions about the relationship between the norm of reciprocity and other socio-moral norms, such as the norms of fairness and welfare.

Here are some of our conjectures. The norms of fairness and welfare together balance behaviors in intergroup contexts. The norm of fairness motivates individuals to initiate positive actions (e.g., distributing resources fairly) toward outgroup members, and also promotes fair reciprocation of positive acts by outgroup members (e.g., rewarding efforts fairly). The norm of welfare inhibits extreme unprovoked aggression toward outgroup members and also limits the severity of escalated retaliation.
against outgroup members. Therefore, these seemingly conflicting socio-moral norms actually complement and balance each other in guiding people’s moral judgments and behaviors.

4. Is infant morality primarily driven by emotions or cognition?

Philosophical debates regarding the emotional versus cognitive basis of morality have caught a lot of attention for several centuries. The emotion-based approach and the principle-based approach hold divergent opinions on whether morality is primarily driven by automatic emotions or by tacit principles (i.e., cognition).

The emotion-based approach emphasizes the role of emotions in the generation and development of moral judgments (e.g., Haidt, 2001; 2008; Haidt & Joseph, 2007). Reasoning is the “slave of the passions” and provides mostly post-hoc justifications for gut reactions. This approach predicts that early moral expectations should primarily be triggered by external emotional stimuli, such as how violent actions are (a bottom-up process).

On the other hand, the principle-based approach argues that moral principles are the foundation upon which moral emotions are based (e.g., Dupoux & Jacob, 2007; Dwyer, 2007; Geraci & Surian, in press; Hamlin, Wynn, Bloom, 2007; Jackendoff, 2007; Olson & Spelke, 2008; Mikhail, 2007; Premack, 2007; Raw, 1971). Infants and toddlers have some rudimentary expectations of moral principles and they can use these moral principles to make sense of individuals’ actions in specific contexts. Their moral emotions, if any, are triggered by a violation of moral principles. According to this approach, early socio-moral expectations are generated by applying internal moral principles to particular contextual factors (a top-down process).

The present projects provide evidence for the principle-based approach. Emotional responses alone cannot explain early socio-moral expectations. The results of our projects demonstrate that children are engaged in cognitive processing as they evaluate particular contextual factors and generate socio-moral expectations accordingly. For example, infants and toddlers can already (1) consider the intergroup context (e.g., whether individuals belong to the same or different groups); (2) consider the prior history of interactions among the individuals (e.g., the presence of prior provocation, the previous actions targeted
toward the individual’s ingroup members as well as those targeted toward the individual him/herself; (3) attribute differentiated mental states to each member of a group (e.g., the individuals’ motivations and intentions; their knowledge and ignorance; and their false beliefs and pretense). Therefore, infants and toddlers can already take into account not only the outcomes of actions, but also the contexts in which the actions occur. They deploy multiple cognitive processes (e.g., feature recognition, social categorization, event memory, psychological reasoning, etc.) and apply moral principles flexibly to novel contexts. This early sensitivity to these contextual factors supports the principle-based (as opposed to the emotion-based) approach.

In addition, the developmental gap between prosocial behavior and prosocial emotions may suggest that the emergence of morality does not rely on moral emotions. Prior research suggests that children’s prosocial behaviors appear before the emergence of prosocial emotions. During the first year of life, object sharing and instrumental helping behaviors already emerge, even though children do not yet show empathic responses toward another’s distress; during the second year of life, indifference toward another’s distress decreases and prosocial behavior is more evident; during the third year of life, prosocial actions and emotions become more regulated and selective such that it is show to some but not to all potential recipients (Hay & Rheingold, 1983; Vaish et al., 2009; Werneken & Tomasello, 2007; Zahn-Waxler et al., 1992; Zahn-Waxler et al., 2001). Therefore, it is unlikely that infant morality is based solely on emotions or relies on gut feelings, since there is a delay in the development of moral emotions.

5. Does morality emerge early in life?

As discussed in Chapter 1, two very broad views have been offered on the origin of morality. On the one hand, the late-acquisition view argues that children acquire morality through extensive learning and socialization. On the other hand, the early-emergence view argues that children are born with a rich innate basis for understanding social interactions and that the knowledge of morality emerges early in life.

The present projects provide evidence that (1) infants and toddlers, as bystanders, expect unfamiliar individuals to reciprocate or retaliate in accordance with the norm of reciprocity; (2) they are already sensitive to novel markers of social groups and expect individuals to favor their group members in
accordance with the norm of ingroup loyalty; (3) their expectations of reciprocity are modulated by considerations of ingroup loyalty in a predictable way; (4) they take contextual factors into account when predicting or interpreting others’ actions and interactions. Therefore, infants and toddlers already have rich socio-moral expectations about the interactions among individuals. These expectations emerge early in life. The results in both infants and toddlers also provide evidence for the continuity of these early socio-moral expectations.

6. How does social experience influence early socio-moral expectations?

The results from the present projects support the principle-based, early-emergence view of early socio-moral expectations. This is not to say, of course, that experience plays little role in the development of morality. The early socio-moral expectations will be consolidated or modified by later experience at least in both the family and culture levels.

First of all, parental practices related to morality vary from family to family. For instance, great diversity exists regarding how to help children cope with intergroup relations and how to teach children strategies for succeeding in mainstream society. For example, different families have different practices dealing with the issue of race. These parental practices can be grouped into four kinds. (1) *Cultural socialization* refers to the parental practices that teach children about their racial heritage and history; they promote cultural customs and traditions, as well as children’s cultural, racial, and ethnic pride. For example, parents may talk to children about important historical or cultural figures; expose children to culturally relevant books, music, and stories; celebrate cultural holidays; eat ethnic foods; and encourage children to use their family’s native language. (2) *Preparation for bias* refers to parental efforts to promote their children’s awareness of discrimination and prepare them to cope with it. Discussions about discrimination often emerge as a theme of parents’ narratives. Some helpful strategies regarding how to deal with discrimination may be offered by the parents. (3) *Promotion of mistrust* refers to practices that emphasize the need for wariness and distrust in interracial interactions. Mistrust may be communicated in parents’ cautions or warnings to children about other racial groups, or in their cautions about barriers to success. Parents may teach their children defensive racial protocols and emphasize social distance or
Egalitarianism and silence refers to the practice that parents either explicitly encourage their children to value individual qualities over racial group membership or avoid any mention of race in discussions with their children (Hughes, et al., 2006).

Certain parents may also emphasize some norms more than others. As mentioned in the beginning of Chapter 1, children are regularly exposed to various moral lessons (e.g., Miller et al., 2001). Parents immerse children in environments that are rich in stories and examples related to morality. Those stories and examples may trigger the innate moral modules and be connected to socio-moral principles through the parent-and child discourses.

The influence of experience on socio-moral development not only occurs in the family level, but also in the culture level. Here are several ways that culture may shape young children’s moral judgment and behavior.

(1) Children learn the culturally-appropriate paths to justice in reciprocity and more specifically retaliation. For example, although cultural practices tend to stick to the norm of reciprocity at a fundamental level, the “payment currency” is extremely variable from culture to culture. The price for misdeeds can be physical punishment, social ridicule, withdrawal of support, or demand for material compensation. Reciprocity interacts with various features of a culture, such as social organization. People living in nomadic hunter-gatherer bands are more likely to choose self-redress revenge for retaliation, whereas people living in societies with complex social organizations are more likely to turn to the police or the government for third-party punishment (e.g., Fry, 2006). Therefore, social environment has an important role in shaping the specific paths that people choose to reciprocate or retaliate.

(2) The norm of ingroup loyalty is enriched and elaborated by social information that children are exposed to in daily life. Although children already pay attention to social group markers during infancy, they still need to acquire further information about what or who constitutes an ingroup or an outgroup member. They also need to learn the specific expressions of ingroup loyalty in their social context, through observation or direct experience (e.g., Hirschfeld, 2001).
Children gradually understand the meanings of particular group markers through cultural learning. For example, Dunham, Baron, and Banaji (2006) examined whether children’s and adults’ implicit attitudes are sensitive to the cultural prestige of a given group. Participants’ implicit attitudes toward both a relatively high-status outgroup and a relatively low-status outgroup were measured by the Implicit Attitude Test (IAT). For European American participants, Japanese have a relatively high status while African Americans have a relatively low status; for Japanese participants, European Americans have a relatively high status while African Americans have a relatively low status. The results were consistent across age groups. They suggest that across both populations and types of outgroup tested, implicit bias toward ingroup members was robustly present by age 6. Moreover, implicit bias against low-status outgroup members emerges early and remains stable with age, while implicit bias against high-status outgroup members appears to undergo a gradual, age-related decline in strength. Therefore, the prestige enjoyed by a group moderates implicit bias as greater knowledge of group status is acquired through cultural learning.

The results of the present projects also support the malleability of early social categorization. Our experimenters naturally varied in terms of color of skin, hair, eyes and so on. However, when they were merely marked by some artificial or transient feature (such as their labels and outfits), infants and toddlers were already able to group them accordingly and form different expectations for ingroup and outgroup members.

(3) Children have to learn the significance of various norms and how they are implemented in a particular culture. Dupoux and Jacob (2007) argue that each culture has its own ranking and elaboration of a finite set of universal moral norms. Work by Shweder and colleagues (1995) expanded on this idea by analyzing the practice of arranging “who sleep by whom” at night among Indian and North-American middle-class families. They then extracted moral values from these practices. They found that both cultures extremely valued incest avoidance. Therefore sexually-active unmarried males and females within the family should avoid co-sleeping, which may lead to either sexual temptation, or suspicion of sexual contact. However, the two cultures had different emphasis for the sleeping arrangements of their
children. Indian families emphasized protection of the vulnerable: children are highly valued members of the family and they are needy and fragile. Therefore, they should not be left alone at night. North American families emphasized autonomy: children are highly valued members of the family and while they are needy and fragile, they should be encouraged to be alone at night so that they can learn to be self-reliant and independent. Children are woven into different cultural practices and learn about the unique ranking and elaboration of various norms.

(4) Children’s socio-moral behavior is also synchronized with historical and societal changes. With modernization and globalization, the scale and complexity of human communities have significantly increased, which may influence people’s socio-moral behavior. Recent research found that the bigger a community is, the more the community members would like to punish norm-violators; the more integrated a community’s market is (as indicated by the percentage of purchased calories), and the more people value fairness (Henrich, et al., 2010). The on-going change of societal complexity may indicate that each time period leaves a historically-distinctive imprint on the socio-moral behavior of children growing up in that period. For example, Chen and his colleagues (2005) found different patterns of relations between shyness and social adjustment in 3 cohorts (1990, 1998, and 2002) of elementary-school children during multiple phases of the societal transition in China. Shyness was first associated with social and academic achievement in the 1990 cohort. Then, the associations became weaker in the 1998 cohort, followed by shyness being associated with peer rejection, school problems, and depression in the 2002 cohort.

Conclusion

Infants and toddlers have rudimentary socio-moral expectations about actions and interactions within social contexts. They are sensitive to various markers of social groups and their expectations of reciprocity are modulated by considerations of ingroup loyalty. Socio-moral expectations emerge early in life and are more likely to be based on a small set of socio-moral principles.
FIGURES

**Figure 1**: Schematic drawing of the events shown during the familiarization and test trials in the ingroup condition of Experiment 1 (with or without provocation).

**Familiarization Trial**

**Test Trial**

**One-Block Event**

**Some-Towers Event**

**One-Block Event with No Provocation**

**Some-Towers Event with No Provocation**
Figure 2: Mean looking times at the one-block and some-towers events in the ingroup condition of Experiment 1 (with or without provocation).
Figure 3: Schematic drawing of the events shown during the familiarization and test trials in the outgroup condition of Experiment 2 (with or without provocation).

**Familiarization Trial**

**Test Trial**

**One-Block Event**

**Some-Towers Event**

**One-Block Event with No Provocation**

**Some-Towers Event with No Provocation**
Figure 4: Mean looking times at the one-block and some-towers events in the outgroup condition of Experiment 2 (with or without provocation).
Figure 5: Schematic drawing of the events shown during the familiarization and test trials in the ingroup and outgroup conditions of Experiment 3 (with provocation).

**Experiment 3**

**Ingroup Condition**

Familiarization Trial (same as in Experiment 1)

Test Trial

All-Towers Event

**Outgroup Condition**

Familiarization Event (same as in Experiment 2)

Test Trial

All-Towers Event
Figure 6: Mean looking times at the one-block, some-towers, and all-towers events in the ingroup and outgroup conditions of Experiments 1 through 3 (with provocation).
Figure 7: Schematic drawing of the events shown during the familiarization and test trials in the outgroup condition of Experiment 4.

Familiarization Trials 1 and 2

Familiarization Trials 3 and 4

Pretest Trial

Test Trial

Hinder Event

Help Event
**Figure 8**: Schematic drawing of the events shown during the familiarization and test trials in the ingroup condition of Experiment 4.

**Familiarization Trials 1 and 2**

![Familiarization Trials 1 and 2]

**Familiarization Trials 3 and 4**

![Familiarization Trials 3 and 4]

**Pretest Trial**

![Pretest Trial]

**Test Trial**

**Hinder Event**

![Hinder Event]

**Help Event**

![Help Event]
Figure 9: Schematic drawing of the events shown during the familiarization and test trials in the outgroup-no-provocation condition of Experiment 4.

Familiarization Trials 1 and 2 (same as in the outgroup condition)

Pretest Trial (same as in the outgroup condition)

Test Trial

Hinder Event (same as in the outgroup condition)

Help Event (same as in the outgroup condition)
Figure 10: Mean looking times at the help and hinder event in the outgroup, outgroup no-provocation, and ingroup conditions of Experiments 4.
**Figure 11**: Schematic drawing of the events shown during the familiarization and test trials in the outgroup condition of Experiment 5 (the presence or absence situation).

**Familiarization Trials 1 and 2** (same as in the outgroup condition of Experiment 4)

**Familiarization Trials 3 and 4** (Presence Situation)

**Familiarization Trials 3 and 4** (Absence Situation)

**Pretest Trial** (same as in the outgroup condition of Experiment 4)

**Test Trial**

**Help Event** (same as in the outgroup condition of Experiment 4)
Figure 12: Schematic drawing of the events shown during the familiarization and test trials in the ingroup condition of Experiment 5 (the presence or absence situation).

Familiarization Trials 1 and 2 (same as in the ingroup condition of Experiment 4)

Familiarization Trials 3 and 4 (Presence Situation)

Familiarization Trials 3 and 4 (Absence Situation)

Pretest Trial (same as in the ingroup condition of Experiment 4)

Test Trial

Help Event (same as in the ingroup condition of Experiment 4)
Figure 13: Mean looking times at the help event in the outgroup and ingroup conditions of Experiments 5 (the presence versus absence situation).
Figure 14: Schematic drawing of the events shown during the familiarization trials in the outgroup condition of Experiment 6.

Familiarization Trial 1

Familiarization Trial 2
Figure 15: Schematic drawing of the events shown during the test trial in the outgroup condition of Experiment 6.

Test Trial

Hinder Event

Help Event
Figure 16: Schematic drawing of the events shown during the familiarization trials in the ingroup condition of Experiment 6.

Familiarization Trial 1

Familiarization Trial 2
Figure 17: Schematic drawing of the events shown during the test trial in the outgroup condition of Experiment 6.
Figure 18: Schematic drawing of the events shown during the familiarization and test trials in the outgroup no-provocation condition of Experiment 6.

Familiarization Trials 1 and 2

Test Trial

Hinder Event (same as in the outgroup condition)

Help Event (same as in the outgroup condition)
Figure 19: Mean looking times in the outgroup, outgroup no-provocation, and ingroup conditions of Experiment 6 and in the all-ingroup condition of Experiment 7.
Figure 20: Schematic drawing of the events shown during the familiarization trials in the all-ingroup condition of Experiment 7.

Familiarization Trial 1

Familiarization Trial 2
Figure 21: Schematic drawing of the events shown during the test trial in the all-ingroup condition of Experiment 7.

Test Trial

Hinder Event (same as in the outgroup condition)

Help Event (same as in the outgroup condition)
Figure 22: Schematic drawing of the events shown during the familiarization and test trials in the ingroup condition of Experiment 8.

Familiarization Trials 1 and 2

Test Trial

Refuse Event

Accept Event

Coordinate Event
Figure 23: Mean looking times at the refuse, accept, and coordinate events in the ingroup condition of Experiment 8 and in the outgroup condition of Experiment 9.
Figure 24: Schematic drawing of the events shown during the familiarization and test trials in the outgroup condition of Experiment 9.

Familiarization Trials 1 and 2

Test Trial

Refuse Event (same as in Experiment 8)

Accept Event (same as in Experiment 8)

Coordinate Event (same as in Experiment 8)
**Figure 25**: Schematic drawing of the events shown during the test trial in the unmarked condition of Experiment 10.

**Test Trial**

**Refuse Event (same as in Experiment 8)**

**Coordinate Event (same as in Experiment 8)**
Figure 26: Mean looking times at the refuse and coordinate events in the ingroup condition of Experiment 8, in the outgroup condition of Experiment 9, and in the unmarked condition of Experiment 10.
Figure 27: An account of early expectations about individuals’ initial, retaliatory, and reciprocatory actions in intergroup contexts.

<table>
<thead>
<tr>
<th></th>
<th>Ingroup Condition</th>
<th>Outgroup Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Actions</td>
<td>extremely moderately mildly neutral mildly moderately extremely negative negative negative positive positive positive positive negative negative negative positive positive positive positive</td>
<td>extremely moderately mildly neutral mildly moderately extremely negative negative negative positive positive positive positive negative negative negative positive positive positive positive</td>
</tr>
<tr>
<td>Retaliatory Actions</td>
<td>extremely moderately mildly neutral mildly moderately extremely negative negative negative positive positive positive positive</td>
<td>extremely moderately mildly neutral mildly moderately extremely negative negative negative positive positive positive positive negative negative negative positive positive positive positive</td>
</tr>
<tr>
<td>Reciprocatory Actions</td>
<td>extremely moderately mildly neutral mildly moderately extremely negative negative negative positive positive positive positive</td>
<td>extremely moderately mildly neutral mildly moderately extremely negative negative negative positive positive positive positive negative negative negative positive positive positive positive</td>
</tr>
</tbody>
</table>
REFERENCES


He, Z., Bolz, M., Baillargeon, R., & Premack, D. (under review). If you steal her cookie, she may tear your sticker: expectation of reciprocity in infants.


Sloane, S., Baillargeon, R., & Premack, D. (under review). Do infants have a sense of fairness?


