Investigating the role of theory of mind in cooperative and competitive behaviors using approaches from cognitive neuroscience and developmental psychology

Lily Tsoi

A dissertation

submitted to the Faculty of

the Department of Psychology

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

June 2018

Investigating the role of theory of mind in cooperative and competitive behaviors using approaches from cognitive neuroscience and developmental psychology

Lily Tsoi

Advisor: Liane L. Young, Ph.D.

People are often quite attuned to the minds around them, but it's unclear whether the tendency to consider the minds of others differs depending on the context. Research on intergroup processes and interpersonal relations reveal that the tendency to consider the minds of others depend on factors like group membership; however, interactions with ingroup members and outgroup members tend to conflate with cooperative interactions and competitive interactions, respectively. Cooperation and competition are two categories of interactions that encompass most of collective human behavior and thus provide natural categories for grouping social behaviors. We test the idea that people's tendencies to consider the minds of others depend on the type of social interaction by primarily focusing on cooperation and competition. Papers 1 and 2 directly compare theory of mind across cooperative and competitive contexts, whereas Paper 3 aims to understand the role of theory of mind in supporting one important aspect of cooperation—a sense of fairness—by studying responses to different forms of unfairness across a spectrum of ages in children. Altogether, these results show an influence of theory of mind on social evaluations and social behaviors and support the idea that sensitivity to context may emerge early in life but becomes more difficult to detect over time.

TABLE OF CONTENTS

Table of Contents ii
List of tables iv
List of figures v
Acknowledgments vi
1.0 Introduction 1
2.0 Paper 1: Developmental evidence for enhanced theory of mind for competitive versus cooperative contexts 6
2.1 Introduction 7
2.2 General Methods 10
2.2.1 Participants 10
2.3 Study 1 11
2.3.1 Methods 11
2.3.2 Results 13
2.3.3 Methods 14
2.3.4 Results 16
2.4 Comparisons between Study 1 and Study 2 19
2.4.1 Methods 19
2.4.2 Results 19
2.5 General Discussion 20
2.6 References 22
2.7 Appendix 46
2.7.1 Script for Study 1 46
2.7.2 Script for Study 2 47
2.8 Supplementary Material 51
2.8.1 Participant Information 51
2.8.2 Study 1: Additional Analyses 54
2.8.3 Study 2: Additional Analyses 58
3.0 Paper 2: The impact of reading about cooperative and competitive actions on brain activity in regions implicated in theory of mind 59
3.1 Introduction 60
3.2 Methods 66
3.2.1 Participants 66
3.2.2 Materials 66
3.2.3 Procedure 67
3.2.4 fMRI acquisition and preprocessing 68

3.2.5 Data Analysis 68
3.3 Results 73
3.3.1 Behavioral analyses 73
3.3.2 Analyses of response magnitudes 75
3.3.3 Analyses of spatial patterns 76
3.3.4 Analyses on the relationship between neural data and behavioral data 78
3.4 Discussion 83
3.5 References 87
3.6 Appendix 113
3.7 Supplementary Material 120
3.7.1 Theory of mind localizer 120

4.0 Paper 3: Individual differences in theory of mind predict inequity aversion in children 121
4.1 Introduction 122
4.2 Methods 127
4.2.1 Participants 127
4.2.2 Design 129
4.2.3 Procedure 129
4.2.4 Coding and Analyses 133
4.3 Results 135
4.3.1 Effects of the theory of mind manipulation 137
4.3.2 Relationship between inequity aversion and individual differences in theory of mind 139
4.3.3 Relationship between inequity aversion and individual differences in empathy 140
4.4 Discussion 142
4.4.1 Limitations and Caveats 147
4.5 Conclusion 148
4.6 References 150
4.7 Appendix 175
4.7.1 Script for the experimental task 175
4.7.2 Reading the Mind in the Eyes Test (child version) 183
4.7.3 Griffith Empathy Measure 213
4.8 Supplementary Material 220
4.8.1 Participants 220
4.8.2 Full model 221
4.8.3 Supplemental results: With age as a continuous variable 223
4.8.4 Supplemental results: Interactions with RMET score 224

5.0 General Discussion 227

6.0 References 232

LIST OF TABLES

Paper 1

None

Paper 2

None

Paper 3

Table 1. Model output from a reduced model that includes all the two-way interactions of interest (Distribution x Condition, Distribution x RMET score, Distribution x Griffith score) (in bold). 136

LIST OF FIGURES

Paper 1

Figure 1. Proportion of responses that are correct broken down by Age Category, Condition, and Question Type. 14

Figure 2. Performance on the game. Proportion of stickers won, (a) broken down by Age Category and Condition, (b) broken down by Condition and Total Number of Possible Stickers for the 4-year-old age group, and (c) depicted as a histogram broken down by Age Category and Condition. 18

Paper 2

Figure 1. Performance (top) and reaction time (bottom) on the True/False question, split by condition. 74

Figure 2. Post-scan ratings of how nice the agent is, split by condition. 75

Figure 3. ROI results. 78

Figure 4. Relationship between PSC and probability of making a rating (as estimated by the model), separated by condition and ROI. 80

Figure 5. Relationship between predicted values for performance on True/False questions (as estimated by the model) and PSC, separated by ROI and Condition. 82

Paper 3

Figure 1. Experimental procedure. 130

Figure 2. Rejections of equal and unequal distributions by Inequity Type (advantageous inequity [top] vs. disadvantageous inequity [bottom]), Condition (theory of mind vs. control), and Decider Age Group (6&7 vs. 8&9). 138

Figure 3. Rejections of equal and unequal distributions by each of the individual difference measures: Reading the Mind in the Eyes Test (top) and Griffith Empathy Measure (bottom). 141

ACKNOWLEDGMENTS

I would like to thank:

- Liane, who has been an amazing mentor and confidante throughout my seven years at BC. I am so grateful to have been a part of her lab family. She is, without a doubt, one of my favorite people ever.
- Katie, for introducing me to developmental psychology and for having me be a part of her awesome lab. I don't know if I'll miss the Skittles, but I will definitely miss the honey pretzel twists.
- Lab mates, past and present, for all their academic and emotional support.
    - Morality Lab: I will miss taking the BC shuttle with Minjae, witnessing BoKyung's acrobatic skills, sharing stouts with Justin, and seeing Josh's notifications to drink water pop up on his computer.
    - Cooperation Lab: I will miss popping into Gorana's office for random visits, eating Turkish snacks with Melisa, and watching Laurent make sushi masterpieces.
- Elizabeth and Josh G. for their feedback and willingness to read this entire dissertation. And thanks for being kind to graduate students stumbling through the academic life.
- Diana for shining that light at the end of the tunnel that is graduate school.

- Bonnie for helping me reinforce the idea that life is a work in progress.
- The Dinner Party, which provided me with a community I never thought I needed.
- My mom for making me bring food back to Boston whenever I went home. Her fears about me starving in grad school was completely unfounded, but the food was always very much appreciated.
- Pauline for being by my side since the fourth grade. Chatting with her over the phone while walking around the Reservoir has been a favorite pastime.
- Eric, my partner through thick and thin. I am really excited to close the distance and start the next chapter of my life with him.

1.0 INTRODUCTION

To successfully navigate the social world, one must have some understanding of how others think (e.g., what they want; what they believe, whether they mean you harm). Theory of mind—the capacity to attribute, infer, and reason about minds—plays a key role in moral judgments of right and wrong (Young, Cushman, Hauser, & Saxe, 2007), in social evaluations of others (Todd & Galinsky, 2014), and in facilitating coordination, cooperation, and communication (Epley & Waytz, 2010). The literature on this topic has largely focused on two main questions: (1) Do X [e.g., infants, children, non-human primates, individuals with autism] have a theory of mind? (2) When do people consider the minds of others? This dissertation will focus on the second question but is largely motivated by research addressing the first question.

When do people consider the minds of others?

People consider the minds of others quite readily and spontaneously. People are quick to consider mental states like intentions when making moral judgments (J. Decety & Cacioppo, 2012; Malle & Holbrook, 2012) and when reading about people's behaviors even when explicit mental state information is not presented (Young & Saxe, 2009). In fact, evidence suggests that people—when presented with any stimuli that can be construed as social—spontaneously use mental state terms to describe the situation (Heider & Simmel, 1944). Moreover, moving shapes perceived as animate (Wheatley, Milleville, & Martin, 2007) and even single-frame pictures containing humans (Wagner, Kelley, & Heatherton, 2011) are sufficient to elicit activity in brain regions that support theory of mind. This tendency to consider the minds of others is so strong that people sometimes even see minds and anthropomorphize non-human entities like pets and gadgets (Epley, Akalis, Waytz, & Cacioppo, 2008; Epley, Waytz, & Cacioppo, 2007; Waytz, Epley, & Cacioppo, 2010; Waytz, Morewedge, et al., 2010a). One thing is clear: people appear to be quite attuned to the minds around them.

However, are people's tendencies to engage in theory of mind constant across different contexts? Research on intergroup processes and interpersonal relations reveal that the tendency to consider the minds of others depend on who those others are (e.g., whether they belong to the same social groups or different social groups; Kelman, 1973; Leyens et al., 2000; Opotow, 1990; Struch & Schwartz, 1989). However, interactions with ingroup members and outgroup members tend to conflate with cooperative interactions and competitive interactions, respectively. Cooperation and competition are two categories of interactions that encompass most of collective human behavior and thus provide natural categories for grouping social behaviors. Perhaps people's tendencies to consider the minds of others depend more on the type of social interaction than the social groups of the interaction partner.

Some research supports the idea that the cooperation / competition distinction may matter for theory of mind. One line of work examines theory of mind capacities in non-human primates, addressing the debate of whether non-human animals are adept at predicting the behaviors of other creatures because they have the ability to understand what is going on other animals' minds or because they have the ability to understand environmental cues that can be used to infer behavior without having to interpret them as content inside others' minds (Lurz, 2011; Premack

& Woodruff, 1978). Earlier non-human primate studies supported the notion that non-human primates have no understanding of mental states (Heyes, 1998; Povinelli, Rulf, & Bierschwale, 1994; Povinelli & Vonk, 2003), but later work revealed an important point to consider: by and large, these tasks were cooperative, but the social lives of most non-human primates mainly revolve around competition (Hare, 2001; Hare & Tomasello, 2004; Lyons & Santos, 2006). When theory of mind tasks were framed in a competitive manner, researchers were successful in showing rudimentary capacities for theory of mind (Flombaum & Santos, 2005; Hare, Call, Agnetta, & Tomasello, 2000; Hare, Call, & Tomasello, 2006; Hare & Tomasello, 2004; Phillips, Barnes, Mahajan, Yamaguchi, & Santos, 2009; Santos, Nissen, & Ferrugia, 2006).

Another line of work examines theory of mind capacities in children and how these capacities are impaired in autism spectrum disorder (ASD), a neurodevelopmental disorder characterized by difficulties with social communication and interaction. One cornerstone of social competence is the ability to understand that people can have beliefs that differ from one's own beliefs and beliefs that differ from reality (false beliefs). The capacity to explicitly represent mental states such as beliefs emerges between three and five years of age (Saxe, Carey, & Kanwisher, 2004; Wellman, Cross, & Watson, 2001; but see Onishi & Baillargeon, 2005). Children younger than this consistently fail tasks that require them to understand that others may hold false beliefs. For instance, if a child puts chocolate in the kitchen cupboard and his mother moves the chocolate to a drawer while the child is away, preschool children will correctly say that the child will search for the chocolate in the cupboard, whereas children younger than this will typically answer incorrectly (Wimmer & Perner, 1983). Preschool and even older children with ASD, however, fail these tasks and are generally found to be impaired in theory of mind relative to neurotypical controls matched in age, language ability, and IQ (Simon Baron-Cohen, 2000; Happé, 1995; Yirmiya, Erel, Shaked, & Solomonica-Levi, 1998). Some researchers, however, argue that poor performance on these tasks can be explained in terms of motivation (Peterson, Slaughter, Peterson, & Premack, 2013). That is, individuals with autism may not be motivated to keep up with a conversation or getting an arbitrary question correct in a laboratory setting, even when they are offered candy for correct responses (Begeer, Rieffe, Terwogt, & Stockmann, 2003). Instead, they may be more motivated to consider others' mental states in more naturalistic and relevant situations such as competition (e.g., fighting with a sibling over a specific toy). Indeed, even when children with ASD are unable to perform well on standard false belief tasks, they are able to perform well on competitive games that require them to engage in theory of mind (Peterson et al., 2013).

Overview of the present research

The overarching goal of this dissertation is to examine whether neurotypical children and adults show a sensitivity to social context when engaging in theory of mind. Papers 1 and 2 directly compare theory of mind across cooperative and competitive contexts, whereas Paper 3 aims to understand the role of theory of mind in supporting one important aspect of cooperation—a sense of fairness—by studying responses to different forms of unfairness. In **Paper 1**, we used a developmental approach to test whether neurotypical children are better able to understand false beliefs arising from competitive interactions than false beliefs arising from cooperative interactions, and whether children are better able to plant false beliefs in another's mind to achieve a competitive goal versus a cooperative goal. We found that children were better at understanding false beliefs in competitive than cooperative contexts and better at applying their understanding in competitive interactions than cooperative interactions. In **Paper 2**, we used functional magnetic resonance imaging (fMRI) in neurotypical adults to examine whether reading about someone broadly cooperating or broadly competing with another person elicits different levels of activity in brain regions implicated in theory of mind. We found no evidence of greater engagement of theory of mind for competition vs. cooperation or vice versa and no evidence that brain regions implicated in theory of mind encode information about cooperation or competition. However, we did find that activity in these regions tracked with accurate belief understanding and subsequent impressions of people in the vignettes and varies depending on the context. In **Paper 3**, we used a developmental approach to examine whether theory of mind plays a role in children's responses to inequity that benefits the self compared to inequity that benefits others. We found that a theory of mind manipulation had no significant effects on children's aversion to both forms of inequity, but that individual differences in theory of mind tracked with greater rejection to both forms of inequity. Altogether, these results show an influence of theory of mind on social evaluations and social behaviors and support the idea that sensitivity to context may emerge early in life but is reduced across time.

2.0 Paper 1: Developmental evidence for enhanced theory of mind for competitive versus cooperative contexts

There is a debate regarding the function of theory of mind (ToM), the capacity to infer, attribute, and reason about mental states. On the one hand are evolutionary and psychological work suggesting that ToM is greater for competition than cooperation. On the other hand are findings and theories promoting greater ToM for cooperation than competition. We investigate the question of whether ToM is greater for competition than cooperation or vice versa by examining the period of development during which explicit ToM comes online. In two studies, we examined preschool children's abilities to explicitly express an understanding of false beliefs—a key marker of ToM—and ability to apply that understanding in first-person social interactions in competitive and cooperative contexts. Our findings reveal that preschool children are better at understanding false beliefs and applying that understanding in competitive contexts than in cooperative contexts.

This paper is co-authored with Kiley Hamlin, Adam Waytz, Andrew Scott Baron, and Liane Young.

2.1 INTRODUCTION

Cooperation and competition comprise two basic forms of social interaction. At first glance, successful cooperative and competitive interactions both appear to require the capacity to infer, attribute, and reason about people's minds (e.g., thoughts, beliefs, intentions), a capacity often referred to as theory of mind (ToM). Indeed, to successfully help another person, one must understand both that the person needs help and what he or she needs help with. Similarly, to compete effectively against another person, one must understand what one's opponent is thinking in order to oppose him or her effectively.

Although ToM facilitates both cooperation and competition, some evidence suggests that ToM has developed primarily in service of competitive aims. Prior work on the evolutionary origins of ToM provides evidence for rudimentary ToM capacities in non-human primates in the ecologically salient domain of competition (e.g., over scarce resources such as food), as compared to cooperation (e.g., Hare, 2001; Hare, Call, Agnetta, & Tomasello, 2000; Hare, Call, & Tomasello, 2006; Hare & Tomasello, 2004; Lyons & Santos, 2006; Melis, Call, & Tomasello, 2006). The primarily competitive nature of social interactions among non-human primates and environmental pressures such as limited resources (e.g., for food and mating opportunities) may have favored individuals who could represent the perceptions and simple beliefs of conspecifics — an ability that may have been preserved in the hominid lineage. These findings are in line with evolutionary accounts of ToM as having evolved for Machiavellian aims (Byrne & Corp, 2004; Byrne & Whiten, 1988). Even among human children and adults, agents that display negative behavior, as compared to neutral or positive behavior, are particularly strong triggers for ToM in the service of understanding those agents' present and future behaviors (Morewedge, 2009; Vaish, Grossmann, & Woodward, 2008; Waytz, Morewedge, et al., 2010b).

On the other hand are findings and theories suggesting that, in humans, ToM may have emerged to facilitate cooperation. Unlike our closest ape relatives, humans are "cooperative breeders": individuals distantly related or unrelated to a child often serve as caregivers ("alloparents"), engaging in active food sharing and providing shelter and protection (Hrdy, 2009). It has been argued that the need to identify individuals in one's environment who are most likely to provide optimal care has facilitated the emergence of advanced ToM in humans. This conclusion is supported by the observation that other cooperative breeders in the primate lineage, most notably callitrichids, are largely deficient to apes and various other non-cooperative breeding primates in general cognitive functioning, but nevertheless have relatively advanced social-cognitive abilities (Burkart, Hrdy, & Van Schaik, 2009; but see Thornton & McAuliffe, 2015). Relatedly, experimental work in human adults reveals a greater tendency for people to consider the minds of ingroup members (who are more likely to be cooperators) than outgroup members (or competitors) (Kelman, 1973; Leyens et al., 2000; Opotow, 1990; Struch & Schwartz, 1989).

We previously investigated the question of whether ToM is greater for competition vs. cooperation or vice versa in adults, using overall levels of activity in brain regions implicated in ToM as a proxy for the cognitive process of ToM (Tsoi, Dungan, Waytz, & Young, 2016). We found that while overall levels of activity within ToM regions were similar for cooperative and competitive interactions, these regions nevertheless encoded information separating cooperation from competition in their spatial patterns of activity. Given these results, it may appear that ToM is neither recruited more for competition than cooperation or vice versa, though the distinction between cooperation and competition remains relevant to social interactions.

The approach we take in the current work is to examine the period of development during which explicit ToM comes online. Many developmental studies of ToM have focused on false belief understanding, that is, the understanding that people can have beliefs that contradict reality (Wimmer & Perner, 1983). The ability to understand false beliefs has been shown to emerge during the preschool years, approximately 4 years of age (for a review, Wellman, Cross, & Watson, 2001). At this point in development, children start being able to make explicit predictions and inferences about people's false beliefs. While there is growing evidence of implicit false belief understanding in infants (e.g., Onishi & Baillargeon, 2005; for a review, see Scott & Baillargeon, 2017), the focus of the current studies is on (1) the explicit expression of false belief understanding and (2) the application of that understanding to first-person social interactions.

Here, in two studies, we test whether children, at the age during which explicit ToM comes online, are better at understanding false beliefs in competitive contexts than cooperative contexts or vice versa. In Study 1, we build on an extensive body of literature on preschool children's explicit performance on a classic false belief task (Josef Perner & Roessler, 2012; Wellman et al., 2001; Wimmer & Perner, 1983) and test whether preschool children's false belief understanding is better for processing mean (competitive) versus nice (cooperative) interactions. In Study 2, we depart from traditional ways of measuring ToM as tapped by Study 1, that is, in third-party contexts during which participants observe and assess the behaviors of others (non-interaction partners). Instead, we test ToM during first-person social interactions and use the capacity for deception (i.e., planting a false belief in someone else's mind) as a proxy for ToM (Chandler, Fritz, & Hala, 1989; Hala, Chandler, & Fritz, 1991; Lee, 2013; Premack & Woodruff, 1978; Woodruff & Premack, 1979). Specifically, we examine whether preschool children are better able to plant a false belief in another's mind to achieve a competitive goal (i.e., to be the sole winner of stickers) versus a cooperative goal (i.e., to be joint winners, together, of stickers). The ability to deceive (i.e., plant false beliefs in another's mind) provides a reliable measure of mental state understanding given that a successful lie requires children to understand both their mental states as well as their interaction partner's mental states.

If competitive vs. cooperative contexts (or vice versa) enhance false belief understanding in children, it may do so in one of two ways: (1) enhancement may only affect children at certain age groups (specifically among four-year-olds whose understanding of false belief is just emerging), or (2) enhancement may affect children across all age groups.

2.2 GENERAL METHODS

2.2.1 Participants

Six hundred and seventy-seven children were recruited from a community-based science center and tested in a soundproof room dedicated to behavioral science research. A legal guardian provided informed consent for all children.

Most participants were assigned to both Studies 1 and 2 (N=627). We had a predetermined goal of 60 participants per cell (roughly 30 per gender), based on the cells in Study 2 and our assumption that dropouts would be high (an outcome typically found in science centers; e.g., Workshop on Research and Museum Partnerships, Cognitive Development Society Meeting, October 2015; Gonzalez, Steele, & Baron, 2017; Gonzalez, Dunlop, & Baron, 2016). We aimed to stop when we estimated that we reached that target. Because the policy at the science center is to not turn away participants if they want to participate, we sometimes went beyond our stopping rule for overpopulated ages. For Study 1, the final sample consisted of 537 participants: 147 three-year-olds (74 females), 266 four-year-olds (137 females), and 124 five-year-olds (55 females). For Study 2, the final sample consisted of 541 participants: 166 three-year-olds (89 females), 251 four-year-olds (130 females), and 124 five-year-olds (54 females). Breakdown of sample size per cell and detailed information on exclusion criteria are reported in Supplementary Material.

2.3 STUDY 1

2.3.1 Methods

Procedure

Participants were introduced to a modified version of the Sally-Anne task (Baron-Cohen, 1985) in the form of a live puppet show. Participants were assigned to either the *Nice Anne* condition or the *Mean Anne* condition (counterbalanced across participants; see complete script in Supplementary Material). In the *Nice Anne* condition, Anne, who is a nice girl, moves Sally's ball from the basket to the closet while Sally is away because she wanted to help Sally. In the *Mean Anne* condition, Anne, who is a mean girl, moves Sally's ball from the basket to the closet while Sally is away because she wanted to trick Sally. After the puppet show, participants are asked the following questions: (1) Where will Sally look?, (2) Where does Sally think her ball is?, (3) Should Anne and Sally be friends?, (4) Is Anne a nice girl or not a nice girl?, (5) Is Sally a nice girl or not a nice girl?. The order in which Questions 1 and 2 were asked was counterbalanced across participants. The focus of this paper is on responses to Questions 1, 2, and 4; descriptive statistics for responses to the remaining Questions 3 and 5 are provided in Supplementary Material.

Analyses

Analyses were conducted in R (version 3.3.3; R Core Team, 2015). Responses were analyzed using a Generalized Linear Mixed Model (GLMM) with binary response terms (correct [1] or incorrect [0]). We were primarily interested in whether responses (correct vs. incorrect) depended on several predictor variables. The main predictor of interest was Condition (Mean Anne or Nice Anne). Because the standard question ("Where will Sally look?") might be difficult in that it requires integrating a belief about Sally's mental state as well as knowledge of how mental states can affect motor behavior, we also included a question probing just the belief ("Where does Sally think her ball is?")—hence, we included Question Type as a predictor (manipulated within-participant). Age Category[1] (three, four, or five), and Gender (male or female) were also included. We examined the three-way interaction between Condition, Question Type, and Age Category, the three two-way interactions (Condition x Question Type, Condition x Age Category, and Question Type x Age Category), and the main effects of these variables. Participant was entered as a random effect. To assess the importance of our predictors of interest, we performed likelihood ratio tests (LRTs) and examined whether the model including a given term provided a significantly better fit to the data than the model without that term.

[1] Analyses using Age as a continuous variable revealed the same pattern of results, reported in Supplementary Material.

2.3.2 Results

Overall, we did not see any interactions involving Condition (Mean Anne or Nice Anne). That is, likelihood ratio tests revealed no three-way interaction between Condition, Question Type, and Age Category ($\chi^2(2) = 3.719$, $p = 0.16$), no two-way interaction between Condition and Age Category ($\chi^2(2) = 0.687$, $p = 0.71$), and no two-way interaction between Condition and Question Type ($\chi^2(1) = 2.344$, $p = 0.126$). There was, however, an interaction between Question Type and Age Category ($\chi^2(2) = 10.866$, $p = 0.004$): the effect of Question Type differed across the three age groups (Figure S2). Pairwise contrasts performed at each age group revealed that the log odds of getting the "Where does Sally think her ball is?" question correct was significantly greater than the log odds of getting the "Where will Sally look?" question correct among 4-year-olds ($z = 4.438$, $p < 0.001$) and 5-year-olds ($z = 3.785$, $p < 0.001$), but not among 3-year-olds ($z = 0.495$, $p = 0.62$).

More critically, we found a significant main effect of Condition ($\chi^2(1) = 7.136$, $p = 0.008$): the log odds of providing a correct response was significantly greater for the *Mean Anne* condition than for the *Nice Anne* condition (Figure 1). Entering age as a continuous variable revealed a similar pattern of results: a significant main effect of Condition (see Supplementary Material). We also restricted our analyses to participants who responded to the question "Is Anne a nice girl or not a nice girl?" in a manner congruent with the condition to which they were assigned; this question served as a comprehension check, but excluding people who did not get this question correct did not change the general pattern of results (Supplementary Material).

Even though the interaction between Condition and Age Category was not significant, we nevertheless performed contrasts examining the difference between conditions at each Age Category and for each Question Type. For the question, "Where does Sally think her ball is?", the difference between *Mean Anne* and *Nice Anne* was significant for 3-year-olds (z = 1.997, p = 0.046) but not for 4-year-olds (z = -0.327, p = 0.74) or for 5-year-olds (z = 0.564, p = 0.57). On the other hand, for the question, "Where will Sally look?", the difference between *Mean Anne* and *Nice Anne* was significant for 4-year-olds (z = 2.568, p = 0.01) but not for 3-year-olds (z = 1.069, p = 0.29) or for 5-year-olds (z = 1.260, p = 0.21). These results reveal that children are more likely to respond correctly to questions about beliefs in competitive contexts than cooperative contexts at age 3 and more difficult questions about beliefs and behavior in competitive contexts than cooperative contexts at age 4.

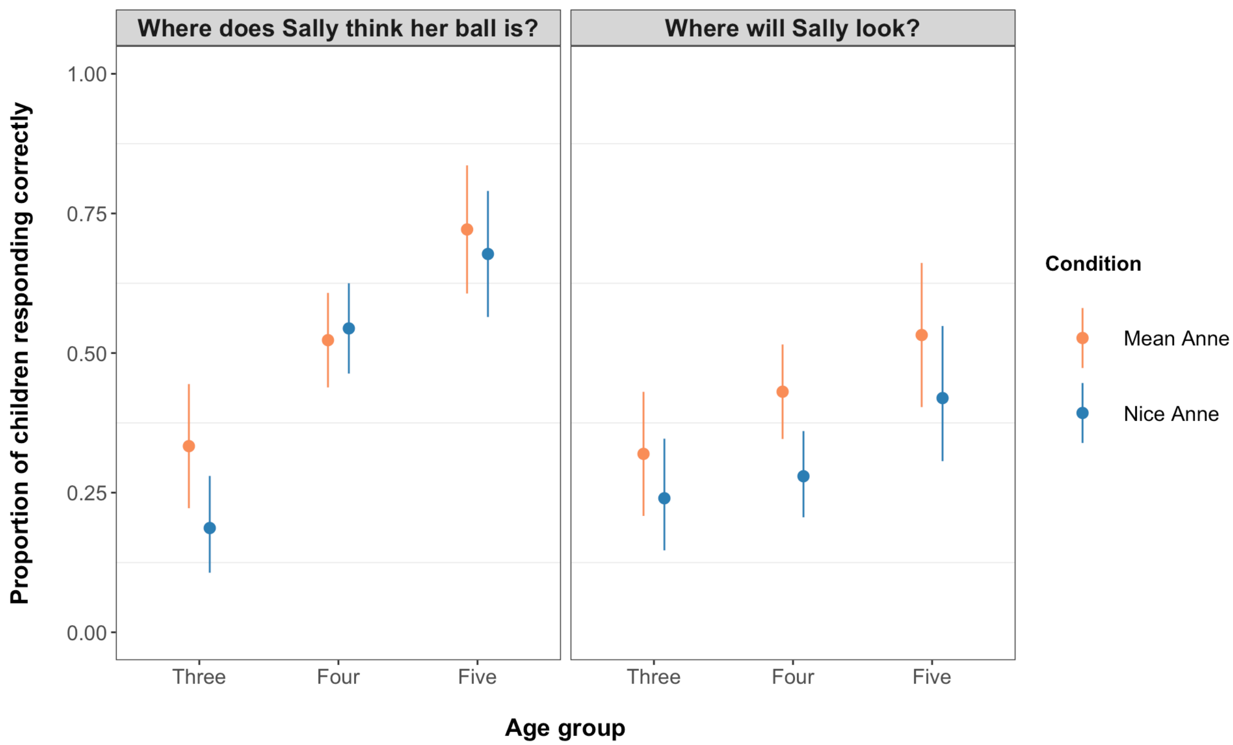


Figure 1. Proportion of responses that are correct broken down by Age Category, Condition, and Question Type. Error bars denote 95% CIs.

2.3.3 Methods

Procedure

In Study 2, participants were introduced to a two-person game involving stickers, where the goal of the game was to get as many stickers as possible. Each participant was assigned to either the *Competition* condition or the *Cooperation* condition (see complete script in Supplementary Material). For both conditions, the participant was instructed to hide two stickers in one of two cups while a second player (a second experimenter, hereafter referred to as E2) had her eyes closed. The participant was instructed to respond however he or she wanted (e.g., by pointing to either Cup #1 or Cup #2) when E2 opened her eyes and asked the participant where the stickers were. E2 would then make a guess as to where the stickers were solely based on the participant's response. In the *Competition* condition, only one person won stickers at a time: if E2 guessed correctly, she kept both stickers, but if she guessed incorrectly, the participant got to keep both stickers. In the *Cooperation* condition, both players could win stickers at the same time: if E2 guessed correctly, neither player got any stickers, but if she guessed incorrectly, the participant and E2 each got to keep one of the two stickers. In order to succeed on either task, the participant would have to plant a false belief in E2's mind. Each participant played four rounds of this game; thus, the participant had the opportunity to win up to 8 stickers in the *Competition* condition and up to 4 stickers in the *Cooperation* condition.

We included two additional conditions for the 4-year-old group to address a potential concern: that a difference in performance between the *Competition* and *Cooperation* conditions could be driven the difference in the maximum number of stickers a person could win (8 stickers versus 4 stickers). Participants may be more motivated to plant a false belief in another's mind when faced with a bigger prize (e.g., more stickers). Thus, we included a condition in which participants had the opportunity to win up to 4 stickers in the *Competition* condition and another in which participants could win up to 8 stickers in the *Cooperation* condition (see scripts in

Supplementary Material). Because we expected any potential difference to emerge around 4 years of age, as predicted by prior work, we recruited only 4-year-olds for these two conditions. For the 3-year-olds, we introduced an additional control condition (*Pompom* condition) to address the concern that poor performance on this task among this age group could be attributed to difficulties with pointing. In this condition, 3-year-olds were instructed to place pompoms in front of one of the two cups (in lieu of pointing to the cup) when responding to E2's question of where the stickers were (see script in Supplementary Material). No difference in performance was found for the pompom versus non-pompom condition ($\chi^2(1) = 0.140$, $p = 0.71$), indicating that the added step of pointing does not negatively affect performance on this task. In our main analyses, we include data collapsing across the *Pompom* and *Non-Pompom* conditions.

Analyses

Analyses were conducted in R (version 3.3.3; R Core Team, 2015). Responses were analyzed using Generalized Linear Models with proportion data. We were primarily interested in whether the proportion of stickers won depended on age and condition; our full model included the following predictor variables: Condition (cooperation or competition), Age Category (three, four, or five), and Gender (male or female). We also examined the two-way interaction between Condition and Age Category. To assess the importance of our predictors of interest, we performed likelihood ratio tests (LRTs) and examined whether the model including a given term provided a significantly better fit to the data than the model without that term.

2.3.4 Results

Analyses revealed no interaction between Condition (cooperation or competition) and Age Category ($\chi^2(2) = 3.5864$, $p = 0.1664$), suggesting that the effect of Condition did not vary across age group. Unsurprisingly, there was a significant main effect of Age Category ($\chi^2(2)$ = 242.25, $p < 0.001$): the log odds of winning stickers was greater among older versus younger age groups. More importantly, and central to the present hypotheses, we found a significant main effect of Condition ($\chi^2(1) = 4.4441$, $p = 0.035$): the log odds of winning stickers was greater for the *Competition* condition than for the *Cooperation* condition. Entering age as a continuous variable revealed the same pattern of results: significant main effects of Condition and Age (see Supplementary Material).

As in Study 1, even though the interaction between Condition and Age Category was not significant, we nevertheless performed contrasts examining the difference between Conditions at each Age Category. These preplanned contrasts revealed a greater difference for competition vs. cooperation among 4-year-olds ($z = 2.489$, $p = 0.0128$), but no difference between cooperation and competition among 3-year-olds ($z = 1.299$, $p = 0.1940$) or 5-year-olds ($z = 0.031$, $p = 0.9753$) (Figure 2a). Together, the results of Study 2 suggest that children may be better at planting false beliefs in others when trying to achieve a competitive goal versus a cooperative goal, with this difference being most pronounced among 4-year-olds.

To examine whether performance in the task among 4-year-olds could be predicted by alternative factors such as the total possible number of stickers a participant could win (4 vs. 8), we examined the effects of Condition and Total Number of Possible Stickers by analyzing just the data with 4-year-olds (Figure 2b). We did not see a significant effect of total possible number of stickers ($\chi^2(1) = 2.4572$, $p = 0.117$), and, more importantly, we still found a similar, albeit marginal, effect of Condition after controlling for Total Number of Possible Stickers ($\chi^2(1)$ = 3.3141, $p = 0.069$).

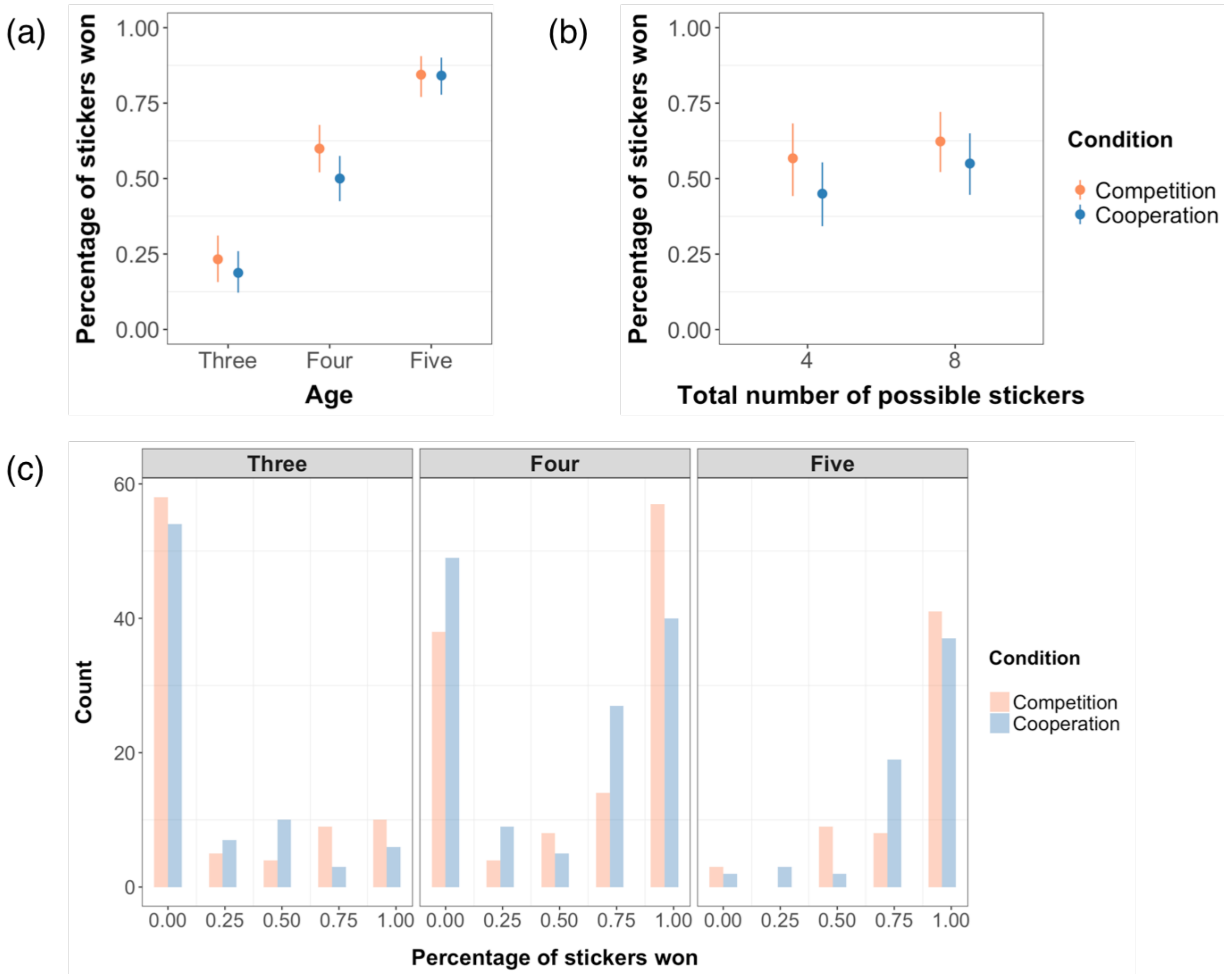


Figure 2. Performance on the game. Proportion of stickers won, (a) broken down by Age Category and Condition, (b) broken down by Condition and Total Number of Possible Stickers for the 4-year-old age group, and (c) depicted as a histogram broken down by Age Category and Condition. Error bars denote 95% CIs.

2.4 COMPARISONS BETWEEN STUDY 1 AND STUDY 2

2.4.1 Methods

Analyses

To examine whether Studies 1 and 2 are tapping into the same ToM construct, we tested whether success in the task for Study 1 could predict success in a different task (Study 2). This is a question we were able to address with the current dataset given that a large subset of our participants completed both studies. Analyses were performed using Generalized Linear Models and only used data from participants who completed both Study 1 and Study 2 ($N = 504$). The response variable was proportion of stickers won in Study 2. Predictor variables included performance in Study 1 (1=correct, 0=incorrect), Condition (Cooperation, Competition), Age Category (3s, 4s, 5s), Gender, the interaction between Condition and Age Category, and the interaction between performance in Study 1 and Age Category.

2.4.2 Results

There was a significant interaction between performance in Study 1 and Age Category ($\chi^2(2) = 13.104$, $p = 0.001$), with contrasts revealing significantly greater log odds of winning stickers in Study 2 for 3- and 4-year olds who answered correctly in Study 1 than for 3- and 4-year olds who answered incorrectly in Study 1 (3s: $z = 5.156$, $p < 0.001$; 4s: $z = 4.187$, $p < 0.001$). The difference among 5-year olds was not significant ($z = 1.200$, $p = 0.23$). These results suggest that performance in Study 1 significantly predicted performance in Study 2 among 3- and 4-year olds but not among 5-year olds.

2.5 GENERAL DISCUSSION

Two studies reveal greater understanding of false beliefs (Study 1) and better application of such understanding (Study 2) for competitive versus cooperative contexts in preschool children. In Study 1, we tested children's false belief understanding using a novel variant of a classic false belief task, tapping into children's abilities to understand people's mental states from the perspective of a third-party observer. In Study 2, we tested children's abilities to plant false beliefs in another's mind, tapping into children's abilities to use mental state information to guide their own behavior. We found a main effect of Condition in both tasks: the log odds of providing a correct response (Study 1) or winning stickers (Study 2) was greater for competitive contexts than for cooperative contexts. In short, we see the basic pattern of results as predicted by evolutionary theories derived from non-human primate work positing better ToM for competition than cooperation.

If ToM evolved to support competitive aims, people may be more attuned to the minds of those they consider to be potential opponents, even if they never actually end up interacting with them. Interestingly, though, in Study 1, participants were not making inferences about Mean Anne; rather, they were making inferences about Sally, who was neither a threat nor an opponent. An alternative possibility is that competitive contexts may trigger ToM such that people are more likely to reason about all minds related to the specific situation. Though our data don't speak directly to this alternative possibility, the results of Study 1 support the idea that children may be more generally attuned to minds in competitive contexts and not specifically to the minds of individual competitors.

What might the observed difference between competition and cooperation reflect? One uninteresting possibility is that the present cooperation conditions make greater processing demands than the competition conditions. That is, better performance in the competition conditions may result from greater difficulty in processing the cooperation conditions. For example, planting a false belief in someone's mind for cooperative reasons (Study 2) might seem unintuitive and unrealistic to our participants. We address this criticism in two ways. First, common activities such as telling white lies or throwing surprise parties also require the ability to plant false beliefs in another's mind for cooperative aims, and indeed, preschool children demonstrate the ability to understand surprises (Hadwin & Perner, 1991; MacLaren & Olson, 1993; Wellman & Banerjee, 1991). Second, this criticism doesn't apply to Study 1, where we nevertheless find condition differences. Moreover, we find consistencies in performance across the two ToM tasks, suggesting that Study 1 and Study 2 tap into the same cognitive capacity. Thus, we believe we are capturing differences between competition and cooperation, rather than superficial task differences.

Whether condition differences remain stable over time is another question. While condition differences appear between ages 3 and 4 for Study 1 and at age 4 for Study 2 (possibly reflecting greater task difficulty for Study 2 as compared to Study 1), no condition difference emerges at age 5 in either study. These results suggest that any specific ToM advantage for competition may emerge when ToM first comes online but then may diminish over time, at least in the context of the present tasks—a suggestion that converges with our prior work showing that brain regions implicated in ToM are recruited to a similar extent for cooperation and competition in adults but these regions nevertheless encode differences relevant to ToM between the two contexts (Tsoi et al., 2016). However, the extent to which this diminished advantage can be attributed to schooling and greater exposure to perspective-taking in cooperative contexts (e.g., sharing with others in school settings) is unclear.

In short, our finding that preschool children show better understanding of false beliefs and are better able to apply that understanding in competitive vs. cooperative contexts contributes to the current debate regarding the function of ToM. Moreover, this work challenges researchers' longstanding treatment of ToM as a unitary process and urges people to treat ToM as more multifaceted.

2.6 REFERENCES

Agresti, A., & Coull, B. A. (1998). Approximate is better than "exact" for interval estimation of binomial proportions. *The American Statistician*, *52*(2), 119–126.

Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Do visual perspective tasks need theory of mind? *NeuroImage*, *30*(3), 1059–1068. https://doi.org/10.1016/j.neuroimage.2005.10.026

Aloise-Young, P. A. (1993). The development of self-presentation: Self-promotion in 6- to 10-year-old children. *Social Cognition*, *11*(2), 201–222. https://doi.org/10.1521/soco.1993.11.2.201

Andreoni, J., & Bernheim, B. D. (2009). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, *77*(5), 1607–1636. https://doi.org/10.3982/ECTA7384

Apfelbaum, E. P., Pauker, K., Ambady, N., Sommers, S. R., & Norton, M. I. (2008). Learning (not) to talk about race: When older children underperform in social categorization. *Developmental Psychology*, *44*(5), 1513–1518. https://doi.org/10.1037/a0012835

Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: A 1-year longitudinal perspective. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00655

Banerjee, R. (2002). Audience effects on self-presentation in childhood. *Social Development*, *11*(4), 487–507. https://doi.org/10.1111/1467-9507.00212

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*(2), 241–251.

Baron-Cohen, Simon. (2000). Theory of mind and autism: A fifteen year review. *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience (2nd Ed.)*, 3–20.

Baron-Cohen, Simon, Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*(1), 37–46.

Baron-Cohen, Simon, Wheelwright, S., Spong, A., Scahill, V., & Lawson, J. (2001). Are intuitive physics and intuitive psychology independent? A test with children with Asperger Syndrome. *Journal of Developmental and Learning Disorders*, *5*(1), 47–78.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015a). Fitting linear mixed-effects models using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015b). Fitting linear mixed-effects models using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Batson, C. D., Thompson, E. R., & Chen, H. (2002). Moral hypocrisy: Addressing some alternatives. *Journal of Personality and Social Psychology*, *83*(2), 330–339. https://doi.org/10.1037/0022-3514.83.2.330

Baumard, N., Mascaro, O., & Chevallier, C. (2012). Preschoolers are able to take merit into account when distributing goods. *Developmental Psychology*, *48*(2), 492–498. https://doi.org/10.1037/a0026598

Begeer, S., Rieffe, C., Terwogt, M. M., & Stockmann, L. (2003). Theory of mind–based action in children from the autism spectrum. *Journal of Autism and Developmental Disorders*, *33*(5), 479–487.

Blake, P. R., McAuliffe, K., Corbit, J., Callaghan, T. C., Barry, O., Bowie, A., … Warneken, F. (2015). The ontogeny of fairness in seven societies. *Nature*, *528*(7581), 258–261. https://doi.org/10.1038/nature15703

Blake, Peter R., & McAuliffe, K. (2011). "I had so much it didn't seem fair": Eight-year-olds reject two forms of inequity. *Cognition*, *120*(2), 215–224. https://doi.org/10.1016/j.cognition.2011.04.006

Blake, Peter R., McAuliffe, K., & Warneken, F. (2014). The developmental origins of fairness: the knowledge–behavior gap. *Trends in Cognitive Sciences*, *18*(11), 559–561. https://doi.org/10.1016/j.tics.2014.08.003

Blake, Peter R., Piovesan, M., Montinari, N., Warneken, F., & Gino, F. (2015). Prosocial norms in the classroom: The role of self-regulation in following norms of giving. *Journal of Economic Behavior & Organization*, *115*, 18–29. https://doi.org/10.1016/j.jebo.2014.10.004

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, *77*(1), B25-31.

Brosnan, S. F., & Bshary, R. (2016). On potential links between inequity aversion and the structure of interactions for the evolution of cooperation. *Behaviour*, *153*(9–11), 1267–1292. https://doi.org/10.1163/1568539X-00003355

Bruneau, E. G., & Saxe, R. (2012). The power of being heard: The benefits of 'perspective-giving' in the context of intergroup conflict. *Journal of Experimental Social Psychology*, *48*(4), 855–866. https://doi.org/10.1016/j.jesp.2012.02.017

Buhrmester, D., Goldfarb, J., & Cantrell, D. (1992). Self-presentation when sharing with friends and nonfriends. *The Journal of Early Adolescence*, *12*(1), 61–79.

Bull, R., Phillips, L. H., & Conway, C. A. (2008). The role of control functions in mentalizing: Dual-task studies of Theory of Mind and executive function. *Cognition*, *107*(2), 663–672. https://doi.org/10.1016/j.cognition.2007.07.015

Burkart, J. M., Hrdy, S. B., & Van Schaik, C. P. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, *18*(5), 175–186. https://doi.org/10.1002/evan.20222

Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language*, *28*(5), 606–637. https://doi.org/10.1111/mila.12036

Byrne, R. W., & Corp, N. (2004). Neocortex size predicts deception rate in primates. *Proceedings of the Royal Society B: Biological Sciences*, *271*(1549), 1693–1699. https://doi.org/10.1098/rspb.2004.2780

Byrne, Richard W., & Whiten, A. (Eds.). (1988). *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford : New York: Clarendon Press ; Oxford University Press.

Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., & Eickhoff, S. B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure and Function*, *217*(4), 783–796. https://doi.org/10.1007/s00429-012-0380-y

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*(4), 1032–1053.

Carlson, Stephanie M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development*, *11*(2), 73–92. https://doi.org/10.1002/icd.298

Carlson, Stephanie M, Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, *87*(4), 299–319. https://doi.org/10.1016/j.jecp.2004.01.002

Carruthers, P. (2016). Two Systems for Mindreading? *Review of Philosophy and Psychology*, *7*(1), 141–162. https://doi.org/10.1007/s13164-015-0259-y

Chandler, M., Fritz, A. S., & Hala, S. (1989). Small-Scale Deceit: Deception as a Marker of Two-, Three-, and Four-Year-Olds' Early Theories of Mind. *Child Development*, *60*(6), 1263. https://doi.org/10.2307/1130919

Christensen, R. H. B. (2015). *ordinal—Regression Models for Ordinal Data*.

Cikara, M., Bruneau, E. G., & Saxe, R. R. (2011). Us and Them: Intergroup Failures of Empathy. *Current Directions in Psychological Science*, *20*(3), 149–153. https://doi.org/10.1177/0963721411408713

Corbit, J., McAuliffe, K., Callaghan, T. C., Blake, P. R., & Warneken, F. (2017). Children's collaboration induces fairness rather than generosity. *Cognition*, *168*, 344–356. https://doi.org/10.1016/j.cognition.2017.07.006

Cortes, B. P., Demoulin, S., Rodriguez, R. T., Rodriguez, A. P., & Leyens, J.-P. (2005). Infrahumanization or Familiarity? Attribution of Uniquely Human Emotions to the Self, the Ingroup, and the Outgroup. *Personality and Social Psychology Bulletin*, *31*(2), 243–253. https://doi.org/10.1177/0146167204271421

Dadds, M. R., Hunter, K., Hawes, D. J., Frost, A. D. J., Vassallo, S., Bunn, P., … Masry, Y. E. (2008). A measure of cognitive and affective empathy in children using parent ratings. *Child Psychiatry and Human Development*, *39*(2), 111–122. https://doi.org/10.1007/s10578-007-0075-4

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, *33*(1), 67–80. https://doi.org/10.1007/s00199-006-0153-z

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. *Nature*, *446*(7137), 794–796. https://doi.org/10.1038/nature05651

Decety, J., & Cacioppo, S. (2012). The speed of morality: A high-density electrical neuroimaging study. *Journal of Neurophysiology*, *108*(11), 3068–3072. https://doi.org/10.1152/jn.00473.2012

Decety, Jean, Jackson, P. L., Sommerville, J. A., Chaminade, T., & Meltzoff, A. N. (2004). The neural bases of cooperation and competition: An fMRI investigation. *NeuroImage*, *23*(2), 744–751. https://doi.org/10.1016/j.neuroimage.2004.05.025

Devine, R. T., & Hughes, C. (2016). Measuring theory of mind across middle childhood: Reliability and validity of the Silent Films and Strange Stories tasks. *Journal of Experimental Child Psychology*, *149*, 23–40. https://doi.org/10.1016/j.jecp.2015.07.011

Devine, R. T., White, N., Ensor, R., & Hughes, C. (2016). Theory of mind in middle childhood: Longitudinal associations with executive function and social competence. *Developmental Psychology*, *52*(5), 758–771. https://doi.org/10.1037/dev0000105

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *NeuroImage*, *55*(2), 705–712. https://doi.org/10.1016/j.neuroimage.2010.12.040

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, *25*(4), 1325–1335. https://doi.org/10.1016/j.neuroimage.2004.12.034

Engelmann, J. M., Herrmann, E., & Tomasello, M. (2012). Five-year olds, but not chimpanzees, attempt to manage their reputations. *PLoS ONE*, *7*(10), e48433. https://doi.org/10.1371/journal.pone.0048433

Engelmann, J. M., & Rapp, D. J. (2018). The influence of reputational concerns on children's prosociality. *Current Opinion in Psychology*, *20*, 92–95. https://doi.org/10.1016/j.copsyc.2017.08.024

Epley, N., Akalis, S., Waytz, A., & Cacioppo, J. T. (2008). Creating Social Connection Through Inferential Reproduction: Loneliness and Perceived Agency in Gadgets, Gods, and Greyhounds. *Psychological Science*, *19*(2), 114–120. https://doi.org/10.1111/j.1467-9280.2008.02056.x

Epley, N., & Waytz, A. (2010). Mind Perception. In *Handbook of Social Psychology*. John Wiley & Sons, Inc. Retrieved from http://dx.doi.org/10.1002/9780470561119.socpsy001014

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868.

FeldmanHall, O., Dalgleish, T., Thompson, R., Evans, D., Schweizer, S., & Mobbs, D. (2012). Differential neural circuitry and self-interest in real *vs* hypothetical moral decisions. *Social Cognitive and Affective Neuroscience*, *7*(7), 743–751. https://doi.org/10.1093/scan/nss069

FeldmanHall, O., Mobbs, D., Evans, D., Hiscox, L., Navrady, L., & Dalgleish, T. (2012). What we say and what we do: The relationship between real and hypothetical moral choices. *Cognition*, *123*(3), 434–441. https://doi.org/10.1016/j.cognition.2012.02.001

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly*, *50*(3), 274–290.

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition*, *57*(2), 109–128.

Flombaum, J. I., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology*, *15*(5), 447–452. https://doi.org/10.1016/j.cub.2004.12.076

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*(4), 708–724.

Galinsky, Adam D., Maddux, W. W., Gilin, D., & White, J. B. (2008). Why it pays to get inside the head of your opponent: The differential effects of perspective taking and empathy in negotiations. *Psychological Science*, *19*(4), 378–384. https://doi.org/10.1111/j.1467-9280.2008.02096.x

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of "theory of mind" in verbal and nonverbal tasks. *Neuropsychologia*, *38*(1), 11–21.

Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental Science*, *14*(5), 1012–1020. https://doi.org/10.1111/j.1467-7687.2011.01048.x

Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J., & Haxby, J. V. (2007). Two takes on the social brain: A comparison of theory of mind tasks. *Journal of Cognitive Neuroscience*, *19*(11), 1803–1814. https://doi.org/10.1162/jocn.2007.19.11.1803

Gonzalez, Antonya M., Steele, J. R., & Baron, A. S. (2017). Reducing children's implicit racial bias through exposure to positive out-group exemplars. *Child Development*, *88*(1), 123–130. https://doi.org/10.1111/cdev.12582

Gonzalez, Antonya Marie, Dunlop, W. L., & Baron, A. S. (2016). Malleability of implicit associations across development. *Developmental Science*. https://doi.org/10.1111/desc.12481

Greene, Joshua D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105–2108.

Greene, Joshua David. (2013). *Moral tribes: Emotion, reason, and the gap between us and them*. New York: The Penguin Press.

Hadwin, J., & Perner, J. (1991). Pleased and surprised: Children's cognitive theory of emotion. *British Journal of Developmental Psychology*, *9*(2), 215–234. https://doi.org/10.1111/j.2044-835X.1991.tb00872.x

Hala, S., Chandler, M., & Fritz, A. S. (1991). Fledgling theories of mind: Deception as a marker of three-year-olds' understanding of false belief. *Child Development*, *62*(1), 83–97. https://doi.org/10.1111/j.1467-8624.1991.tb01516.x

Hamann, K., Bender, J., & Tomasello, M. (2014). Meritocratic sharing is based on collaboration in 3-year-olds. *Developmental Psychology*, *50*(1), 121–128. https://doi.org/10.1037/a0032965

Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature*, *476*(7360), 328–331. https://doi.org/10.1038/nature10278

Happé, F. G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, *66*(3), 843–855.

Hare, B. (2001). Can competitive paradigms increase the validity of experiments on primate social cognition? *Animal Cognition*, *4*(3–4), 269–280. https://doi.org/10.1007/s100710100084

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*(4), 771–785. https://doi.org/10.1006/anbe.1999.1377

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, *61*(1), 139–151. https://doi.org/10.1006/anbe.2000.1518

Hare, B., Call, J., & Tomasello, M. (2006). Chimpanzees deceive a human competitor by hiding. *Cognition*, *101*(3), 495–514. https://doi.org/10.1016/j.cognition.2005.01.011

Hare, B., & Tomasello, M. (2004). Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. *Animal Behaviour*, *68*(3), 571–581. https://doi.org/10.1016/j.anbehav.2003.11.011

Harris, L. T., & Fiske, S. T. (2007). Social groups that elicit disgust are differentially processed in mPFC. *Social Cognitive and Affective Neuroscience*, *2*(1), 45–51. https://doi.org/10.1093/scan/nsl037

Harris, L. T., & Fiske, S. T. (2009). Social neuroscience evidence for dehumanised perception. *European Review of Social Psychology*, *20*(1), 192–231. https://doi.org/10.1080/10463280902954988

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, *10*(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4

Haslam, N., & Loughnan, S. (2014). Dehumanization and infrahumanization. *Annual Review of Psychology*, *65*(1), 399–423. https://doi.org/10.1146/annurev-psych-010213-115045

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, *57*(2), 243. https://doi.org/10.2307/1416950

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Heyes, C. M. (1998). Theory of mind in nonhuman primates. *The Behavioral and Brain Sciences*, *21*(1), 101–114; discussion 115-148.

Hrdy, S. B. (2009). *Mothers and others: The evolutionary origins of mutual understanding*. Cambridge, Mass: Belknap Press of Harvard University Press.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, *16*(2), 233–253. https://doi.org/10.1111/j.2044-835X.1998.tb00921.x

Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development*, *13*(4), 590–619.

Jackson, J. W. (1993). Realistic group conflict theory: A review and evaluation of the theoretical and empirical literature. *The Psychological Record*, *43*(3), 395–413.

Jenkins, A. C., & Mitchell, J. P. (2010). Mentalizing under uncertainty: Dissociated neural responses to ambiguous and unambiguous mental state inferences. *Cerebral Cortex*, *20*(2), 404–410. https://doi.org/10.1093/cercor/bhp109

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Kahneman, D. (2012). *Thinking, fast and slow*. London: Penguin Books.

Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition*, *109*(2), 224–234. https://doi.org/10.1016/j.cognition.2008.08.010

Kanngiesser, P., & Warneken, F. (2012). Young children consider merit when sharing resources with others. *PLoS ONE*, *7*(8), e43979. https://doi.org/10.1371/journal.pone.0043979

Karg, K., Schmelz, M., Call, J., & Tomasello, M. (2015). The goggles experiment: can chimpanzees use self-experience to infer what a competitor can see? *Animal Behaviour*, *105*, 211–221. https://doi.org/10.1016/j.anbehav.2015.04.028

Kelman, H. G. (1973). Violence without moral restraint: Reflections on the dehumanization of victims and victimizers. *Journal of Social Issues*, *29*(4), 25–61. https://doi.org/10.1111/j.1540-4560.1973.tb00102.x

Koster-Hale, J., Saxe, R., Dungan, J., & Young, L. L. (2013). Decoding moral judgments from neural representations of intentions. *Proceedings of the National Academy of Sciences*, *110*(14), 5648–5653. https://doi.org/10.1073/pnas.1207992110

Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology*, *90*(4), 543–555. https://doi.org/10.1037/0022-3514.90.4.543

Lee, K. (2013). Little Liars: Development of Verbal Deception in Children. *Child Development Perspectives*, *7*(2), 91–96. https://doi.org/10.1111/cdep.12023

Leimgruber, K. L., Shaw, A., Santos, L. R., & Olson, K. R. (2012). Young children are more generous when others are aware of their actions. *PLoS ONE*, *7*(10), e48292. https://doi.org/10.1371/journal.pone.0048292

Leyens, J.-P., Paladino, P. M., Rodriguez-Torres, R., Vaes, J., Demoulin, S., Rodriguez-Perez, A., & Gaunt, R. (2000). The emotional side of prejudice: The attribution of secondary emotions to ingroups and outgroups. *Personality and Social Psychology Review*, *4*(2), 186–197. https://doi.org/10.1207/S15327957PSPR0402_06

Lissek, S., Peters, S., Fuchs, N., Witthaus, H., Nicolas, V., Tegenthoff, M., … Brüne, M. (2008). Cooperation and deception recruit different subsets of the theory-of-mind network. *PLoS ONE*, *3*(4), e2023. https://doi.org/10.1371/journal.pone.0002023

Liu, B., Huang, Z., Xu, G., Jin, Y., Chen, Y., Li, X., … Jing, J. (2016). Altruistic sharing behavior in children: Role of theory of mind and inhibitory control. *Journal of Experimental Child Psychology*, *141*, 222–228. https://doi.org/10.1016/j.jecp.2015.09.010

Loewenstein, G. F., Thompson, L., & Bazerman, M. H. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social Psychology*, *57*(3), 426.

Lurz, R. W. (2011). *Mindreading animals: The debate over what animals know about other minds*. Cambridge, Mass: MIT Press.

Lyons, D. E., & Santos, L. R. (2006). Ecology, domain specificity, and the origins of theory of mind: Is competition the catalyst? *Philosophy Compass*, *1*(5), 481–492. https://doi.org/10.1111/j.1747-9991.2006.00032.x

MacLaren, R., & Olson, D. (1993). Trick or treat: Children's understanding of surprise. *Cognitive Development*, *8*(1), 27–46. https://doi.org/10.1016/0885-2014(93)90003-N

Malle, B. F., & Holbrook, J. (2012). Is there a hierarchy of social inferences? The likelihood and speed of inferring intentionality, mind, and personality. *Journal of Personality and Social Psychology*, *102*(4), 661–684. https://doi.org/10.1037/a0026790

McAuliffe, K., Blake, P. R., Kim, G., Wrangham, R. W., & Warneken, F. (2013). Social influences on inequity aversion in children. *PLoS ONE*, *8*(12), e80966. https://doi.org/10.1371/journal.pone.0080966

McAuliffe, K., Blake, P. R., Steinbeis, N., & Warneken, F. (2017). The developmental foundations of human fairness. *Nature Human Behaviour*, *1*, 0042.

McAuliffe, K., & Dunham, Y. (2016). Group bias in cooperative norm enforcement. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150073. https://doi.org/10.1098/rstb.2015.0073

McPherson Frantz, C., & Janoff-Bulman, R. (2000). Considering Both Sides: The Limits of Perspective Taking. *Basic and Applied Social Psychology*, *22*(1), 31–42. https://doi.org/10.1207/S15324834BASP2201_4

Melis, A. P., Call, J., & Tomasello, M. (2006). Chimpanzees (Pan troglodytes) conceal visual and auditory information from others. *Journal of Comparative Psychology*, *120*(2), 154–162. https://doi.org/10.1037/0735-7036.120.2.154

Morewedge, C. K. (2009). Negativity bias in attribution of external agency. *Journal of Experimental Psychology: General*, *138*(4), 535–545. https://doi.org/10.1037/a0016796

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science (New York, N.Y.)*, *308*(5719), 255–258. https://doi.org/10.1126/science.1107621

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. *Frontiers in Neuroinformatics*, *10*. https://doi.org/10.3389/fninf.2016.00027

Opotow, S. (1990). Moral exclusion and injustice: An introduction. *Journal of Social Issues*, *46*(1), 1–20. https://doi.org/10.1111/j.1540-4560.1990.tb00268.x

Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology*, *1*(1), 33–60.

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental Psychology*, *43*(4), 974–990. https://doi.org/10.1037/0012-1649.43.4.974

Perner, J, & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences*, *3*(9), 337–344.

Perner, Josef, & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences*, *16*(10), 519–525. https://doi.org/10.1016/j.tics.2012.08.004

Peterson, C. C., Slaughter, V., Peterson, J., & Premack, D. (2013). Children with autism can track others' beliefs in a competitive game. *Developmental Science*, *16*(3), 443–450. https://doi.org/10.1111/desc.12040

Phillips, W., Barnes, J. L., Mahajan, N., Yamaguchi, M., & Santos, L. R. (2009). 'Unwilling' versus 'unable': capuchin monkeys' (*Cebus apella*) understanding of human intentional action: Unwilling vs. unable in capuchin monkeys. *Developmental Science*, *12*(6), 938–945. https://doi.org/10.1111/j.1467-7687.2009.00840.x

Povinelli, D. J., Rulf, A. B., & Bierschwale, D. T. (1994). Absence of knowledge attribution and self-recognition in young chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *108*(1), 74–80. https://doi.org/10.1037/0735-7036.108.1.74

Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: suspiciously human? *Trends in Cognitive Sciences*, *7*(4), 157–160. https://doi.org/10.1016/S1364-6613(03)00053-6

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(04), 515. https://doi.org/10.1017/S0140525X00076512

Raihani, N. J., & McAuliffe, K. (2012). Human punishment is motivated by inequity aversion, not a desire for reciprocity. *Biology Letters*, *8*(5), 802–804. https://doi.org/10.1098/rsbl.2012.0470

Reis, H. T., & Gruzen, J. (1976). On mediating equity, equality, and self-interest: The role of self-presentation in social exchange. *Journal of Experimental Social Psychology*, *12*(5), 487–503. https://doi.org/10.1016/0022-1031(76)90079-2

Rizzo, M. T., & Killen, M. (2016). Children's understanding of equity in the context of inequality. *British Journal of Developmental Psychology*, *34*(4), 569–581. https://doi.org/10.1111/bjdp.12150

Rochat, P., Dias, M. D. G., Guo Liping, Broesch, T., Passos-Ferreira, C., Winning, A., & Berg, B. (2009). Fairness in distributive justice by 3- and 5-year-olds across seven cultures. *Journal of Cross-Cultural Psychology*, *40*(3), 416–442. https://doi.org/10.1177/0022022109332844

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and U.S.

preschoolers. *Psychological Science*, *17*(1), 74–81. https://doi.org/10.1111/j.1467-9280.2005.01667.x

Santiesteban, I., Banissy, M. J., Catmur, C., & Bird, G. (2015). Functional lateralization of temporoparietal junction - imitation inhibition, visual perspective-taking and theory of mind. *European Journal of Neuroscience*, *42*(8), 2527–2533. https://doi.org/10.1111/ejn.13036

Santos, L. R., Nissen, A. G., & Ferrugia, J. A. (2006). Rhesus monkeys, Macaca mulatta, know what others can and cannot hear. *Animal Behaviour*, *71*(5), 1175–1181. https://doi.org/10.1016/j.anbehav.2005.10.007

Saxe, R, & Kanwisher, N. (2003). People thinking about thinking peopleThe role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*(4), 1835–1842. https://doi.org/10.1016/S1053-8119(03)00230-1

Saxe, R, & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, *43*(10), 1391–1399. https://doi.org/10.1016/j.neuropsychologia.2005.02.013

Saxe, Rebecca, Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, *55*, 87–124. https://doi.org/10.1146/annurev.psych.55.090902.142044

Saxe, Rebecca, & Powell, L. J. (2006). It's the thought that counts: Specific brain regions for one component of theory of mind. *Psychological Science*, *17*(8), 692–699. https://doi.org/10.1111/j.1467-9280.2006.01768.x

Schmelz, M., Call, J., & Tomasello, M. (2011). Chimpanzees know that others make inferences. *Proceedings of the National Academy of Sciences*, *108*(7), 3077–3079. https://doi.org/10.1073/pnas.1000469108

Schmelz, Martin, & Call, J. (2016). The psychology of primate cooperation and competition: a call for realigning research agendas. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150067. https://doi.org/10.1098/rstb.2015.0067

Schmidt, M. F. H., & Sommerville, J. A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. *PLoS ONE*, *6*(10), e23223. https://doi.org/10.1371/journal.pone.0023223

Schug, J., Takagishi, H., Benech, C., & Okada, H. (2016). The development of theory of mind and positive and negative reciprocity in preschool children. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00888

Schurz, M., Aichhorn, M., Martin, A., & Perner, J. (2013). Common brain areas engaged in false belief reasoning and visual perspective taking: a meta-analysis of functional brain imaging studies. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00712

Scott, R. M., & Baillargeon, R. (2017). Early False-Belief Understanding. *Trends in Cognitive Sciences*, *21*(4), 237–249. https://doi.org/10.1016/j.tics.2017.01.012

Shaw, A., Montinari, N., Piovesan, M., Olson, K. R., Gino, F., & Norton, M. I. (2014). Children develop a veil of fairness. *Journal of Experimental Psychology: General*, *143*(1), 363–375. https://doi.org/10.1037/a0031247

Sheskin, M., Bloom, P., & Wynn, K. (2014). Anti-equality: Social comparison in young children. *Cognition*, *130*(2), 152–156. https://doi.org/10.1016/j.cognition.2013.10.008

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience & Biobehavioral Reviews*, *30*(6), 855–863. https://doi.org/10.1016/j.neubiorev.2006.06.011

Singer, T. (2009). Understanding others: Brain mechanisms of theory of mind and empathy. *Neuroeconomics: Decision Making and the Brain*, 251–268.

Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness? *Psychological Science*, *23*(2), 196–204. https://doi.org/10.1177/0956797611422072

Smith, C. E., Blake, P. R., & Harris, P. L. (2013). I should but I won't: Why young children endorse norms of fair sharing but do not follow them. *PloS One*, *8*(3), e59510. https://doi.org/10.1371/journal.pone.0059510

Struch, N., & Schwartz, S. H. (1989). Intergroup aggression: Its predictors and distinctness from in-group bias. *Journal of Personality and Social Psychology*, *56*(3), 364–373.

Takagishi, H., Koizumi, M., Fujii, T., Schug, J., Kameshima, S., & Yamagishi, T. (2014). The role of cognitive and emotional perspective taking in economic decision making in the Ultimatum Game. *PLoS ONE*, *9*(9), e108462. https://doi.org/10.1371/journal.pone.0108462

Thornton, A., & McAuliffe, K. (2015). Cognitive consequences of cooperative breeding? A critical appraisal: Cognitive consequences of cooperative breeding. *Journal of Zoology*, *295*(1), 12–22. https://doi.org/10.1111/jzo.12198

Todd, A. R., & Galinsky, A. D. (2014). Perspective-Taking as a Strategy for Improving Intergroup Relations: Evidence, Mechanisms, and Qualifications: Perspective-Taking and Intergroup Relations. *Social and Personality Psychology Compass*, *8*(7), 374–387. https://doi.org/10.1111/spc3.12116

Tsoi, L., Dungan, J. A., Chakroff, A., & Young, L. L. (2018). Neural substrates for moral judgments of psychological versus physical harm. *Social Cognitive and Affective Neuroscience*. https://doi.org/10.1093/scan/nsy029

Tsoi, L., Dungan, J., Waytz, A., & Young, L. (2016). Distinct neural patterns of social cognition for cooperation versus competition. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2016.04.069

Tsoi, L., Hamlin, J. K., Baron, A. S., Waytz, A., & Young, L. (under review). Developmental evidence for enhanced theory of mind for competitive versus cooperative contexts.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*(3), 383–403. https://doi.org/10.1037/0033-2909.134.3.383

Valdesolo, P., & DeSteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion*, *11*(2), 262–266. https://doi.org/10.1037/a0021302

Wagner, D. D., Kelley, W. M., & Heatherton, T. F. (2011). Individual differences in the spontaneous recruitment of brain regions supporting mental state understanding when viewing natural social scenes. *Cerebral Cortex*, *21*(12), 2788–2796. https://doi.org/10.1093/cercor/bhr074

Waytz, A., Epley, N., & Cacioppo, J. T. (2010). Social Cognition Unbound: Insights Into Anthropomorphism and Dehumanization. *Current Directions in Psychological Science*, *19*(1), 58–62. https://doi.org/10.1177/0963721409359302

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, *14*(8), 383–388. https://doi.org/10.1016/j.tics.2010.05.006

Waytz, A., Morewedge, C. K., Epley, N., Monteleone, G., Gao, J.-H., & Cacioppo, J. T. (2010a). Making sense by making sentient: Effectance motivation increases anthropomorphism. *Journal of Personality and Social Psychology*, *99*(3), 410–435. https://doi.org/10.1037/a0020240

Waytz, A., Morewedge, C. K., Epley, N., Monteleone, G., Gao, J.-H., & Cacioppo, J. T. (2010b). Making sense by making sentient: Effectance motivation increases anthropomorphism. *Journal of Personality and Social Psychology*, *99*(3), 410–435. https://doi.org/10.1037/a0020240

Waytz, A., & Young, L. (2014). Two motivations for two dimensions of mind. *Journal of Experimental Social Psychology*, *55*, 278–283. https://doi.org/10.1016/j.jesp.2014.08.001

Wellman, H. M., & Banerjee, M. (1991). Mind and emotion: Children's understanding of the emotional consequences of beliefs and desires. *British Journal of Developmental Psychology*, *9*(2), 191–214. https://doi.org/10.1111/j.2044-835X.1991.tb00871.x

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655–684.

Wheatley, T., Milleville, S. C., & Martin, A. (2007). Understanding animate agents: Distinct roles for the social network and mirror system. *Psychological Science*, *18*(6), 469–474. https://doi.org/10.1111/j.1467-9280.2007.01923.x

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*(5), 297–333. https://doi.org/10.1037/h0040934

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and Cooperation. *Psychological Science*, *20*(1), 1–5. https://doi.org/10.1111/j.1467-9280.2008.02253.x

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128.

Woodruff, G., & Premack, D. (1979). Intentional communication in the chimpanzee: The development of deception. *Cognition*, *7*(4), 333–362. https://doi.org/10.1016/0010-0277(79)90021-0

Yirmiya, N., Erel, O., Shaked, M., & Solomonica-Levi, D. (1998). Meta-analyses comparing theory of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. *Psychological Bulletin*, *124*(3), 283–307.

Yoder, K. J., & Decety, J. (2014). The Good, the Bad, and the Just: Justice sensitivity predicts neural response during moral evaluation of actions performed by others. *Journal of Neuroscience*, *34*(12), 4161–4166. https://doi.org/10.1523/JNEUROSCI.4648-13.2014

Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences*, *104*(20), 8235–8240. https://doi.org/10.1073/pnas.0701408104

Young, Liane, & Saxe, R. (2009a). An fMRI investigation of spontaneous mental state inference for moral judgment. *Journal of Cognitive Neuroscience*, *21*(7), 1396–1405. https://doi.org/10.1162/jocn.2009.21137

Young, Liane, & Saxe, R. (2009b). Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia*, *47*(10), 2065–2072. https://doi.org/10.1016/j.neuropsychologia.2009.03.020

Zaki, J., & Ochsner, K. (2016). Empathy. In *Handbook of Emotion* (4th ed., pp. 871–884). New York: Guilford.

2.7 APPENDIX

2.7.1 Script for Study 1

***Nice Anne* Condition**

Hi (child's name)! Let me introduce you to some people.

This is Sally (Sally waves), and this is Anne (Anne waves). Sally is playing with her favourite ball (act playing with ball) when her mom calls her out to lunch. Sally has to go, so she runs off, dropping her ball in the basket on the way out (act dropping ball in basket).

Now Anne sees where Sally put her favourite ball and knows that the ball is supposed to go in the closet, and NOT in the basket. Anne thinks that Sally must have been in such a hurry that she put her ball in the wrong place! Anne is a very nice girl, so she wants to help Sally by moving the ball to the closet. Anne goes and gets the ball out of the basket, and to help Sally, puts it away in the closet. That was very nice. Now Sally comes back after lunch and wants to find her ball.

Where will Sally look?
Where does Sally think her ball is?
Should Anne and Sally be friends?
Is Anne a nice girl or not a nice girl?
Is Sally a nice girl or not a nice girl?

***Mean Anne* Condition**

Hi (child's name)! Let me introduce you to some people.

This is Sally (Sally waves), and this is Anne (Anne waves). Sally is playing with her favourite ball (act playing with ball) when her mom calls her out to lunch. Sally has to go, so she runs off, dropping her ball in the basket on the way out (act dropping ball in basket).

Now Anne sees where Sally put her favourite ball and knows that Sally LOVES to play with it. Anne thinks that Sally must have been in such a hurry that she will definitely come back after lunch to play with her ball again. Anne is not a very nice girl, so she wants to trick Sally by moving the ball to the closet. Anne goes and gets the ball out of the basket, and to trick Sally, puts it away in the closet. That was not very nice. Now Sally comes back after lunch and wants to find her ball.

Where will Sally look?
Where does Sally think her ball is?
Should Anne and Sally be friends?

Is Anne a nice girl or not a nice girl?
Is Sally a nice girl or not a nice girl?

2.7.2 Script for Study 2

***Competition* Condition – 8 stickers**

You like stickers right? Well, E2 likes stickers too. Let me tell you about this game. Ok (child's name), here is the game. The goal of the game is to get as MANY stickers as you each can. You try and get as many as you can, and she'll try and get as many as she can. You and E2 both like these stickers, but in this game, only ONE person can win stickers at a time.

To play the game, I'll ask E2 to close her eyes, and then you can hide two stickers in one of these cups. You can put the stickers in this cup, in this cup or in this cup, in any cup you like, but both stickers have to go in the same cup! Now, after you hide the stickers while E2 can't see you, I will tell E2 she can open her eyes, and she will ask you a question about where the stickers are, and you can tell her whatever you want to. Then E2 will guess where the stickers are.

Now here's how the game goes. If E2 guesses wrong, then YOU get to keep both the stickers, but if E2 guesses right, then SHE gets to keep both the stickers! Get it? If SHE finds the stickers, SHE gets to keep them, but is she DOESN'T find the stickers, than YOU get to keep them!
So we'll play the game the say way every time- you always get to hide the stickers and E2 will ask you a question and then guess where they are. Ok, just to make sure you understand the rules. If E2 finds the stickers, who gets to keep them? If she doesn't find the stickers, who gets to keep them?

Great, ok. Remember, the goal of the game is to win as many stickers as you can! You can take home all the stickers you win.

E1 to E2: Ok E2, close your eyes! No peeking!
E1 to child: ok, hide the stickers! Did you hide them?
E1 to E2: Ok E2, you can open your eyes and ask (child's name) a question!
E2 to child: ok (child's name), now I'm going to ask you, can you show me where the stickers are?
If stickers are in the cup, E1: "there are the stickers!" and gives them to E2. "This time the stickers go to E2!"
If the stickers are not in the cup, E1: "Huh, no stickers", find the stickers and give stickers to the child, "This time the stickers go to you!"
E1: Great! Ok, remember, we want to have as many stickers to take home at the end of the game!

***Cooperation* Condition – 4 stickers**

You like stickers right? Well, E2 likes stickers too. Let me tell you about this game. Ok (child's name), here is the game. The goal of the game is to get as MANY stickers as you each can. You try and get as many as you can, and she'll try and get as many as she can. You and E2 both like these stickers, and in this game, you can BOTH get stickers at the same time.

To play the game, I'll ask E2 to close her eyes, and then you can hide two stickers in one of these cups. You can put the stickers in this cup, in this cup or in this cup, in any cup you like, but both stickers have to go in the same cup! Now, after you hide the stickers while E2 can't see you, I will tell E2 she can open her eyes, and she will ask you a question about where the stickers are, and you can tell her whatever you want to. Then E2 will guess where the stickers are.

Now here's how the game goes. If E2 guesses wrong, then each of you gets to keep one sticker! But if E2 guesses right, then NEITHER of you gets any stickers and they go back in the box! Get it? If SHE finds the stickers, no one gets ANY stickers, but if she DOESN'T find the stickers, then you EACH get to keep one!

So we'll play the game the say way every time- you always get to hide the stickers and E2 will ask you a question and then guess where they are. Ok, just to make sure you understand the rules. If E2 finds the stickers, who gets to keep them? If she doesn't find the stickers, who gets to keep them?

Great, ok. Remember, the goal of the game is to win as many stickers as you can! You can take home all the stickers you win.

E1 to E2: Ok E2, close your eyes! No peeking!
E1 to child: ok, hide the stickers! Did you hide them?
E1 to E2: Ok E2, you can open your eyes and ask (child's name) a question!
E2 to child: ok (child's name), now I'm going to ask you, can you show me where the stickers are?
If stickers are in the cup, E1: "there are the stickers! This time no one gets any stickers!"
If the stickers are not in the cup, E1 "Huh, no stickers", find the stickers and give stickers to the child and E2, "This time you both get a sticker!"
E1: Great! Ok, remember, we want to have as many stickers to take home at the end of the game!

***Cooperation-Sheet* condition for 4-year-olds – 8 stickers**
In the *Cooperation* condition, participants could win at most 4 stickers, whereas in the *Competition* condition, participants could win at most 8 stickers. To address the concern that the number of total stickers that a participant could win may influence how well participants performed in the game, we added a different version of the *Cooperation* condition in which the total number of stickers one could possibly win matches the total number of stickers one could possibly win in the *Competition* condition (8 stickers).

Do you like stickers? Can I please get you to pick out 8 sticker sheets from this box?

You like stickers right? Well, E2 likes stickers too. Let me tell you about this game. Ok (child's name), here is the game. The goal of the game is to get as MANY sheets of stickers as you each can. You try and get as many as you can, and she'll try and get as many as she can. You and E2 both like these sheets of stickers, and in this game, you can BOTH get sticker sheets at the same time.

To play the game, I'll ask E2 to close her eyes, and then you can hide two sheets of stickers in one of these cups. You can put the sticker sheets in this cup, in this cup or in this cup, in any cup you like, but both sticker sheets have to go in the same cup! Now, after you hide the sticker sheets while E2 can't see you, I will tell E2 she can open her eyes, and she will ask you a question about where the stickers sheets are, and you can tell her whatever you want to. Then E2 will guess where the stickers sheets are.

Now here's how the game goes. If E2 guesses wrong, then each of you gets to keep one sticker sheet! But if E2 guesses right, then NEITHER of you gets any sticker sheets and they go back in the box! Get it? If SHE finds the sticker sheets, no one gets ANY sticker sheets, but if she DOESN'T find the sticker sheets, then you EACH get to keep one!

So we'll play the game the say way every time- you always get to hide the sticker sheets and E2 will ask you a question and then guess where they are. Ok, just to make sure you understand the rules. If E2 finds the sticker sheets, who gets to keep them? If she doesn't find the sticker sheets, who gets to keep them?

Great, ok. Remember, the goal of the game is to win as many sticker sheets as you can! You can take home all the sticker sheets you win.

E1 to E2: Ok E2, close your eyes! No peeking!
E1 to child: ok, hide the stickers sheets! Did you hide them?
E1 to E2: Ok E2, you can open your eyes and ask (child's name) a question!
E2 to child: ok (child's name), now I'm going to ask you, can you show me where the sticker sheets are?

If stickers are in the cup, E1: "there are the stickers! This time no one gets any sticker sheets!" If the stickers are not in the cup, E1 "Huh, no sticker sheets", find the stickers and give stickers to the child and E2, "This time you both get a sticker sheet!"

***Competition-Four Stickers* condition for 4-year-olds – 4 stickers**
In the *Cooperation* condition, participants could win at most 4 stickers, whereas in the *Competition* condition, participants could win at most 8 stickers. To address the concern that the number of total stickers that a participant could win may influence how well participants performed in the game, we added a different version of the *Competition* condition in which the total number of stickers one could possibly win matches the total number of stickers one could possibly win in the *Cooperation* condition (4 stickers).

The script was the same as the one for *Competition* condition – 8 stickers, but instead of facing the possibility of earning 2 stickers per round, participants faced the possibility of earning 1 sticker per round.

***Pompom* condition for 3-year-olds**

Because we were concerned that difficulties with pointing among 3-year-olds could have affected their performance on the task, we introduced a control condition (*Pompom* condition) to a subset of 3-year-olds. Instead of pointing to the cup of their choosing, participants in this condition placed a pompom in front of the cup.

E1 to E2: Ok E2, close your eyes! No peeking!
E1 to child: ok, hide the stickers! Did you hide them?
E1 to E2: Ok E2, you can open your eyes and ask (child's name) a question!
E2 to child: ok (child's name), now I'm going to ask you, can you show me where the stickers are by placing the pompom in front of the cup where the stickers are?
If stickers are in the cup, E1: "there are the stickers! This time no one gets any stickers!"
If the stickers are not in the cup, E1 "Huh, no stickers", find the stickers and give stickers to the child and E2, "This time you both get a sticker!"
E1: Great! Ok, remember, we want to have as many stickers to take home at the end of the game!

2.8 SUPPLEMENTARY MATERIAL

2.8.1 Participant Information

We recruited a total of 677 participants. Most participants were assigned to both Studies 1 and 2, and our predetermined goal of 60 participants per cell (roughly 30 per gender) was based on the cells in Study 2.

Study 1:

The final sample consisted of 537 participants (Table S1). Of the 643 participants that were recruited for Study 1, 106 were excluded due to: participant's age being outside our age range of interest (31), incompletion of the task or declining to do the task (12), insufficient understanding of English (10), parental/other interference (9), being a part of experimenter pilot/training (8), lack of attention to the task (6), experimenter error (4), not understanding the task (4), having previously seen or completed the task (3), improper consent forms (2), having a developmental disorder (2), fussing out (2), or having data noted by the experimenter as unusable (12).

Table S1. Final sample breakdown by age, gender, and condition for Study 1

	Female	Male	Total
Mean Anne			
three	37	35	**72**
four	66	64	**130**
five	26	36	**62**
Nice Anne			
three	37	38	**75**
four	71	65	**136**
five	29	33	**62**
Grand Total	**266**	**271**	**537**

Study 2:

The final sample consisted of 541 participants (Table S2). Of the 662 participants that were recruited for Study 2, 121 were excluded due to: age outside our age range of interest (30), incompletion of the task or declining to do the task (21), parental/other interference (11), insufficient understanding of English (9), lack of understanding of the task (9), lack of attention to the task (6), experimenter error (6), indications of not following task instructions (5), participant having previously seen or completed the task (3), invalid consent forms (2), participant having a developmental disorder (2), fussing out (2), being a part of experimenter pilot/training (2), or having data noted by the experimenter as unusable (13).

Table S2. Final sample breakdown by age, gender, and condition for Study 2

	Female	Male	Total
Competition			
three	48	38	**86**
four – 4 stickers	26	26	**52**
four – 8 stickers	39	30	**69**
five	27	34	**61**
Cooperation			
three	41	39	**80**
four – 4 stickers	35	30	**65**
four – 8 stickers	30	35	**65**
five	27	36	**63**
Grand Total	**273**	**268**	**541**

2.8.2 Study 1: Additional Analyses

Descriptive statistics for responses to questions in Study 1

In addition to the two main questions, “Where will Sally look?” and “Where does Sally think her ball is?”, we asked, “Should Anne and Sally be friends?”, “Is Anne a nice girl or not a nice girl?”, and “Is Sally a nice girl or not a nice girl?”. Proportion data for the latter three questions are depicted below (Figure S1).

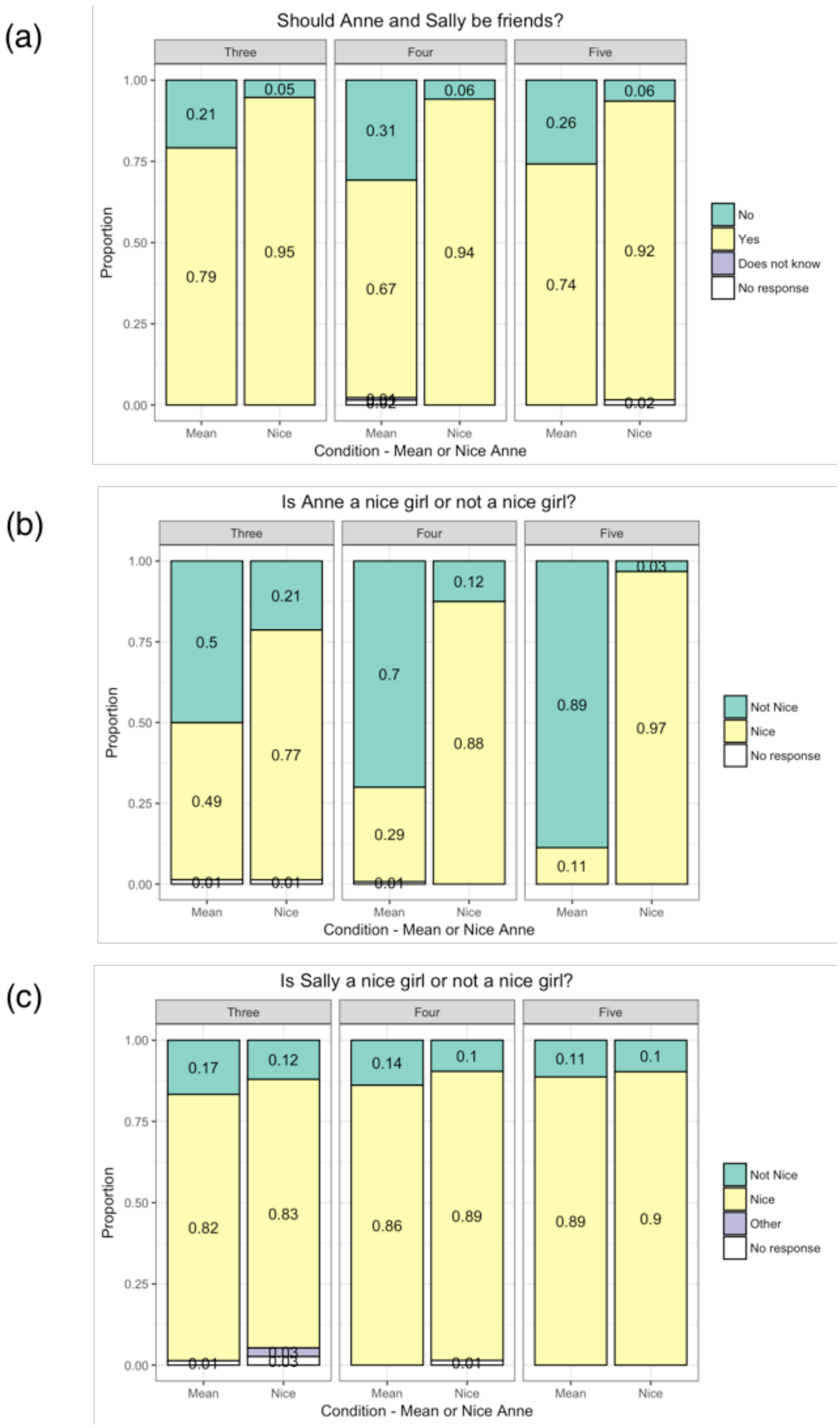


Figure S3. Proportion of participants by response to auxiliary questions in Study 1, broken down by Condition and Age Category.

Breakdown of performance by Age Category and Question Type

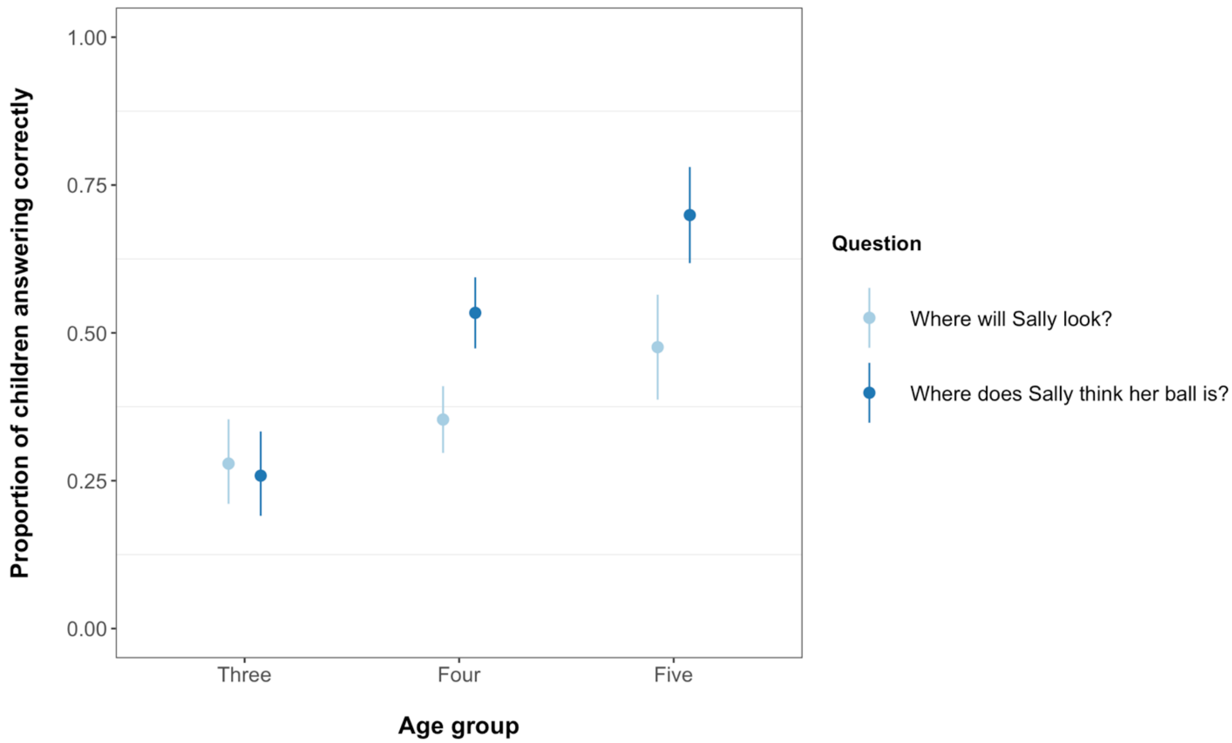


Figure S2. Proportion of participants answering correctly by age group and question. Results are reported in the main text. Bars denote 95% CI.

Analyses limited to participants who responded in a congruent manner

The following analyses restricted the data to those who responded to the question "Is Anne a nice girl or not a nice girl?" in a manner congruent to the condition to which they were assigned. This question acted as a comprehension check, but excluding people who did not get this question correct did not change the general pattern of results. Results were similar to those found in the main text with all usable data. Likelihood ratio tests revealed no significant three-way interaction between Condition, Question Type, and Age Category ($\chi^2(2) = 2.0941$, $p = 0.351$), no significant two-way interaction between Condition and Age Category ($\chi^2(2) = 2.0691$, $p = 0.3554$), a marginally significant two-way interaction between Condition and Question Type ($\chi^2(1) = 3.1647$, p

= 0.075), and a marginally significant interaction between Question Type and Age Category ($\chi^2(2) = 5.4554$, $p = 0.065$).

More importantly for our hypotheses, the main effect of Condition remained significant ($\chi^2(1) = 13.717$, $p < 0.001$): the log odds of providing a correct response was significantly greater for the *Mean Anne* condition than for the *Nice Anne* condition. Entering age as a continuous variable provided a similar pattern of results, revealing, once again, a significant main effect of Condition (see Supplementary section on Analyses with age as a continuous variable).

Analyses with age as a continuous variable

Because some models failed to converge, we dropped participant as a random effect from analyses using age as a continuous variable and analyzed the data using Generalized Linear Models instead of Generalized Linear Mixed Models.

With all usable data (data reported in the main text):

Entering age as a continuous variable did not affect the pattern of results found in the main text. Similar to results reported in the main text, likelihood ratio tests revealed no significant three-way interaction between Condition, Question Type, and Age ($\chi^2(1) = 1.0116$, $p = 0.3145$), no significant two-way interaction between Condition and Age ($\chi^2(1) = 0.682$, $p = 0.409$), and no significant two-way interaction between Condition and Question Type ($\chi^2(1) = 2.3335$, $p = 0.1266$). There was, again, a significant interaction between Question Type and Age Category ($\chi^2(1) = 8.8752$, $p = 0.003$): the effect of Question Type differed by age. Importantly, the main effect of Condition remained significant ($\chi^2(1) = 7.2893$, $p = 0.007$): the log odds of providing a correct response was significantly greater for the *Mean Anne* condition than for the *Nice Anne* condition.

With data limited to participants responding in a congruent manner:

Similar to results reported in the main text, likelihood ratio tests revealed no significant three-way interaction between Condition, Question Type, and Age ($\chi^2(1) = 0.73388$, $p = 0.3916$) and no significant two-way interaction between Condition and Age ($\chi^2(1) = 1.3816$, $p = 0.2398$). There was a marginally significant two-way interaction between Condition and Question Type ($\chi^2(1) = 3.1597$, $p = 0.07548$). There was, again, a significant interaction between Question Type and Age ($\chi^2(1) = 4.4642$, $p = 0.0346$): the effect of Question Type differed by age. Importantly, the main effect of Condition remained significant ($\chi^2(1) = 13.605$, $p < 0.001$): the log odds of providing a correct response was significantly greater for the *Mean Anne* condition than for the *Nice Anne* condition.

2.8.3 Study 2: Additional Analyses

Analyses with age as a continuous variable

Likelihood ratio tests revealed a significant main effect of Context ($\chi^2(1) = 6.2352$, $p = 0.013$) and a significant main effect of the Age category ($\chi^2(1) = 233.08$, $p < 0.001$): that is, the odds of receiving stickers was greater for the *Competition* condition than for the *Cooperation* condition, and the odds of receiving stickers increased with increasing age. This pattern of results is the same found for models with age entered as a categorical variable (see main text).

3.0 PAPER 2: The impact of reading about cooperative and competitive actions on brain activity in regions implicated in theory of mind

Cooperation and competition are two fundamental social contexts that encompasses much of collective human behavior. We investigated the extent to which adults engaged in theory of mind (ToM) across cooperation and competition. Using functional magnetic resonance imaging (fMRI), we scanned participants while they read vignettes of agents impacting someone for malicious / competitive reasons, helpful / cooperative reasons, and non-specified reasons. In the task, they answered questions that required them to reason about the false beliefs of the person impacted by the agent. Results provide no evidence that adults engage in ToM to a greater extent for competition than cooperation or vice versa; furthermore, regions implicated in ToM do not appear to encode the context in which the interaction took place (cooperative, competitive, neutral). However, activity in these regions tracked with post-scan behavioral impression ratings of the agents in the vignettes and performance on the experimental task, though these relationships differed across conditions and brain regions. Together, this work support the idea that while overall levels of activity within regions implicated in ToM do not differ across social contexts, activity in these regions track with accurate belief understanding and subsequent impressions of others and vary depending on the context.

This paper is co-authored with Liane Young.

3.1 INTRODUCTION

People can readily and spontaneously consider the minds of others. For instance, people are quick to think about others' mental states (e.g., desires, beliefs, intentions) when making moral judgments (Decety & Cacioppo, 2012; Malle & Holbrook, 2012) and when reading about people's behaviors even when explicit mental state information is not presented (Young & Saxe, 2009a). Research reveals that this tendency sometimes goes haywire, leading people to see minds in anthropomorphized non-human entities like pets and malfunctioning electronic devices (Epley et al., 2008, 2007; Waytz, Epley, et al., 2010; Waytz, Morewedge, et al., 2010a). The ease with which people engage in 'theory of mind' (ToM) or infer, attribute, and reason about mental states suggests that people may be naturally attuned to the minds around them. However, there may be reason to believe that the extent to which people engage in ToM differs across social contexts. We focus primarily on cooperation and competition, as these two are fundamental social contexts that widely encompass much of collective human behavior. Below are three lines of evidence making distinct predictions regarding ToM across cooperation and competition.

One line of evidence predicts greater ToM for cooperation than for competition. Psychological research on intergroup processes reveals that people consider the minds of others differently depending on factors like group membership (Kelman, 1973; Leyens et al., 2000; Opotow, 1990; Struch & Schwartz, 1989). For instance, people tend to make more attributions of secondary emotions like affection, pride, and nostalgia to ingroup members than to outgroup members, even when ruling out potential alternative explanations like increased familiarity for ingroup members than for outgroup members

(Leyens et al., 2000). Indeed, ingroup members may be viewed as possessing more human essence than outgroup members (Cortes, Demoulin, Rodriguez, Rodriguez, & Leyens, 2005). In a similar vein, people are generally more willing to consider the perspective of people they like in the context of interpersonal conflict (McPherson Frantz & Janoff-Bulman, 2000) and are more likely to attribute greater mental capacities like planning, setting goals, thinking, and remembering to people they like (Kozak, Marsh, & Wegner, 2006). In contrast, people are less likely to consider the mental and emotional experience of outgroup members (Cikara, Bruneau, & Saxe, 2011), and theories such as realistic group conflict theory (Jackson, 1993) suggest that intergroup conflict and hostility could be driven by actual or even perceived competition over limited resources like food, money, or jobs. In fact, evaluating others who are extremely dissimilar to the self fails to elicit activity in regions implicated in ToM (Harris & Fiske, 2007, 2009). The denial of humanness—in the mild form of infrahumanization or the more extreme form of dehumanization—can be found not just in times of war and genocide but also in everyday life. Indeed, infrahumanization can be observed both explicitly and implicitly across groups that differ by ethnicity, race, gender, and disability, just to name a few (Haslam, 2006; Haslam & Loughnan, 2014). Given that ingroup members are more likely to be people with whom we would cooperate (McAuliffe & Dunham, 2016) and outgroup members are more likely to be people with whom we would compete, we may expect to find greater ToM for cooperation than for competition.

A second line of evidence predicts greater ToM for competition than for cooperation. This line of work has focused primarily on rudimentary ToM capacities in non-human primates, inspired by the question of whether ToM is a uniquely human capacity (Premack & Woodruff, 1978). In the 1980s and 1990s, researchers consistently failed to find evidence of ToM in non-human primates, but research in the 2000s revealed the importance of social context in experimental paradigms (Hare, 2001; Hare & Tomasello, 2004). It turned out that early experimental tasks were largely cooperative in nature; when experimental tasks became competitive (e.g., with chimpanzees outsmarting dominant conspecifics for food), chimpanzees displayed rudimentary ToM capacities. Specifically, this work reveals that, in competitive settings, chimpanzees are able to see what others can see and cannot see (Hare, Call, & Tomasello, 2001; Melis, Call, & Tomasello, 2006), what others can hear and cannot hear (Melis et al., 2006), and what others know and don't know (but not what others believe and don't believe) (Kaminski, Call, & Tomasello, 2008). Moreover, they can even make inferences about what others would do based on their own preferences or experience (Karg et al., 2015; Schmelz et al., 2011). One reason for this asymmetry could be due to the fact that social interactions among chimpanzees and many other non-human primates are by and large competitive, leading to the idea that competition may be a more ecologically salient context among these species (Lyons & Santos, 2006; but see Schmelz & Call, 2016). Selective pressures such as limited availability of food may have favored individuals with some capacity to represent the simple beliefs and perceptions of other creatures, and this selective advantage may have persisted in humans. In humans, we see some evidence of this. Children at the age during which they develop explicit ToM in the form of false belief understanding are not only better able to understanding false beliefs in competitive contexts than cooperative contexts, they are also better able to apply that understanding in social interactions to achieve a competitive goal versus a cooperative goal (Tsoi et al., under revision; Paper 1 of this dissertation). Moreover, both children and adults have a greater tendency to use negative behavior to make inferences about agents when compared to positive or neutral behavior (Morewedge, 2009; Vaish et al., 2008; Waytz, Morewedge, et al., 2010a); attempts to gain control over a negative environment or negative agent may drive this asymmetry (Peeters & Czapinski, 1990).

A third line of research predicts no difference in ToM processing for cooperation and competition. Instead, this line of work predicts that both cooperative and competitive interactions will elicit ToM but the motivations behind engagement of ToM may differ. In cooperative contexts, people may be motivated to affiliate and create social bonds with others, whereas in competitive contexts, people may be motivated to predict opponents' actions and gain mastery over their environment (Epley et al., 2007; Waytz, Gray, Epley, & Wegner, 2010; White, 1959). Extensions of this work has revealed that agentic and affiliative motivations both lead people to focus on mental states but different types of mental states (Waytz & Young, 2014).

Given evidence supporting the idea that ToM is greater for competition than cooperation in children and human's closest living relatives, one main question is whether this ToM asymmetry is also present in human adulthood. Functional magnetic resonance imaging (fMRI) provides a promising method for studying this question in adults. One reason is that this method circumvents the need to have participants explicitly state their thoughts or make any self-reports, which could be unreliable and subject to experimenter demand. Moreover, decades of research have revealed a set of regions that are reliably and robustly recruited when people make inferences about mental states as compared to other different types of states (e.g., physical states) (Fletcher et al., 1995;

Gallagher et al., 2000; Gobbini, Koralek, Bryan, Montgomery, & Haxby, 2007; R Saxe & Kanwisher, 2003). This ToM network consists of right and left temporoparietal junction (RTPJ, LTPJ), precuneus (PC), and dorsomedial prefrontal cortex (dmPFC). Because so much is known about the links between ToM and this set of regions, one common way of assessing ToM is to measure overall response magnitudes in these regions as a proxy for the cognitive process of ToM.

This method has been used previously in this lab to examine the exact question of whether ToM is greater for competition than cooperation or vice versa. In the study (Tsoi, Dungan, Waytz, & Young, 2016), adult participants played a cooperative and competitive game with whom they believed was one another person, limiting potential confounds related to liking or familiarity. They found that the mean response magnitude in each ToM region of interest did not differ across cooperation and competition, suggesting that ToM is not recruited more for competition than cooperation or vice versa. What they did find, however, is that all these ToM regions were nevertheless sensitive to the distinction separating cooperative trials from competitive trials; this sensitivity was detected using multivariate pattern analyses (MVPA) to examine spatial patterns of neural activity within ToM regions.

To what extent is this sensitivity about the social context? That is, can this sensitivity be found in the absence of personally-relevant social interactions? The present study examines this question using fMRI. Instead of playing simple economic games, participants in the present study read vignettes in which an agent was either trying to impact another person negatively (broadly competitive), positively (broadly cooperative) or in a non-valenced manner (neutral). These vignettes were variations of the Sally-Anne task used to examine false belief understanding in children (Simon Baron-Cohen, Leslie, & Frith, 1985); more specifically, we modeled the vignettes after the cooperative and competitive scenarios used in a previous study (Tsoi, Hamlin, Baron, Waytz, & Young, under revision; Study 1 of Paper 1 of the dissertation). That is, in each vignette, participants read that an agent did something (e.g., put the ball in the closet), in the hopes of disrupting or annoying the protagonist (competition condition), helping the protagonist (cooperation condition), or something unrelated to the protagonist (neutral condition).

Several predictions can be made from this study. If adults engage in ToM to a similar extent for cooperation and competition, we may see no difference in overall response magnitudes in ToM across condition. If adults engage in ToM differently across contexts, we would observe that difference in overall response magnitudes in ToM ROIs. If regions implicated in ToM were sensitive to information separating cooperative and competitive contexts, we would expect to find above-chance classification levels in ToM ROIs using multivariate pattern analyses. However, if this sensitivity is specific to social interactions that one is personally engaged in, we would not expect to find above-chance classification levels in ToM ROIs in this particular study.

Moreover, in this study, people made post-scan ratings of their impression of each agent in the story (how nice the agent was). In addition to examining whether there were neural differences across conditions, we examined, in an exploratory manner, possible relationships between neural data and behavioral data, with a focus on how neural data predicts subsequent impressions of an agent.

3.2 METHODS

3.2.1 Participants

Thirty-four participants between the ages of 18 and 35 (Mean ± SD: 25.18 ± 4.68; 15 females) were recruited from the Boston area. This sample size was determined by a power analysis using NeuroPower (aiming for 80% power with Random Field Theory error rate control) on a related dataset (Tsoi et al., 2016). All participants were native English speakers, had normal or normal-to-corrected vision, and reported no history of psychiatric or neurological disorders. Participants gave written informed consent and were paid for their participation. The study was approved by the Boston College Institutional Review Board.

3.2.2 Materials

Experimental task

The experimental task contained 30 scenarios (see Appendix). There were three versions of each scenario, corresponding to the three conditions (cooperation, competition, neutral). Each scenario contained three segments: background (10 s), mental state (4 s), and question (6 s). The amount of words in the mental state segment (when information relevant to the condition was presented) was the same across the three conditions (M = 12.33 words). Moreover, items in each condition were pre-tested such that agents in cooperative scenarios were rated as nicer than agents in neutral scenarios, which were rated as nicer than agents in competitive scenarios. Additionally, when we pre-tested, we found no difference in performance on the True/False questions across the three conditions.

ToM localizer task

The ToM localizer task (Dodell-Feder, Koster-Hale, Bedny, & Saxe, 2011; R Saxe & Kanwisher, 2003) was used to functionally define the following regions implicated in ToM: the rTPJ, lTPJ, precuneus, and dmPFC. More details can be found in Supplementary Material.

3.2.3 Procedure

In the experimental task, participants were scanned while reading 10 scenarios in each of three conditions (cooperation, competition, neutral) for a total of 30 unique scenarios (see Appendix). During the question segment of each scenario, participants answered a True / False question using a response box. There was a 12 s fixation period between trials. Six scenarios were presented in each 3.2-minute run for a total of 5 runs. Each scenario had three versions corresponding to the three conditions; participants were only presented with one version for each scenario, and this was counterbalanced across participants.

Participants also completed two runs of the ToM localizer task. After the scan session, participants filled out a post-scan survey in which they rated the extent to which they thought the agent was nice on a scale from 1 [not at all] to 7 [very]. They also filled out a demographics questionnaire.

3.2.4 fMRI acquisition and preprocessing

fMRI data were collected using a 32-channel head coil in a 3T MAGNETOM Prisma scanner at the Athinoula A. Martinos Center at the McGovern Institute for Brain Research at the Massachusetts Institute of Technology. Functional data were collected as thirty-two near-axial slices (3-mm isotropic voxels; 0.59-mm gap) acquired in an interleaved fashion using a gradient echo planar imaging (EPI) sequence (TR=2s, TE=30ms, FA=90°; FOV = 210 mm x 210 mm). Anatomical data were collected with a T1-weighted multi-echo magnetization prepared rapid acquisition echo image sequence (TR=2530 ms, TE=1.69 ms; FA=7°; 1-mm isotropic voxels; FOV=256 mm x 256 mm).

Data preprocessing and analyses were performed using SPM12 and custom software. Functional data were corrected for slice timing, realigned to the first EPI, spatially normalized onto a common brain space (Montreal Neurological Institute; MNI), spatially smoothed using a Gaussian filter (full-width half-maximum = 8-mm kernel), and high-pass filtered (128 s).

3.2.5 Data Analysis

We provide an outline of our analysis plan: first, we conducted analyses of in-scanner responses to examine whether the percentage correct or reaction time differed across conditions (cooperation vs. competition vs. neutral). We also conducted analyses of post-scan ratings to examine whether participants rated the agent differently across conditions. We then conducted whole-brain and ROI univariate analyses to examine response magnitudes across conditions. These two analyses served different functions:

whole-brain analyses were conducted to reveal involvement of any region in the brain in processing actions that were cooperative vs. competitive. ROI analyses, on the other hand, allowed us to examine neural activity across conditions within regions implicated in social cognition as defined using an independent functional localizer task; these analyses allowed us to directly test our hypotheses regarding the role of ToM in processing cooperative and competitive actions. Notably, the ROI analyses were conducted by taking into account by-participant and by-item variance, allowing us to make inferences that can generalize past the specific sample of participants tested as well as the specific items we used in the present study (Judd, Westfall, & Kenny, 2012). We also conducted whole-brain and ROI multivariate analyses to examine whether the cooperation/competition dimension of social actions is a feature encoded in spatial patterns of activity across the brain and specifically within ToM regions, respectively. As a conservative measure, we focused on activity during the mental state and question segments since the background segment did not contain any information about condition. Additionally, we conducted analyses examining the relationship between neural data and post-scan ratings as well as in-scanner behavioral performance.

Behavioral analyses

We were interested in seeing whether performance and ratings differed across conditions. Behavioral analyses were conducted in R (version 3.3.3). Responses to True / False questions were analyzed using a generalized linear mixed model (glmm) with a binary response term (correct / incorrect); the model was run using the package 'lme4' (Bates, Mächler, Bolker, & Walker, 2015a). Reaction times were analyzed using linear mixed models and were also run using the package 'lme4'. Ratings were analyzed using cumulative link mixed models (clmm) with an ordinal response term (from a scale of 1 to 7); the model was run using the package 'ordinal' (Christensen, 2015). Models included condition as a predictor, and participant, item, and trial order as random effects. To assess the importance of our predictor of interest, we performed a likelihood ratio test (LRT) to test whether the model including the predictor of interest would provide a better fit to the data than a model without that term.

Analyses of response magnitudes

Whole-brain analyses. Preprocessed images were analyzed using a general linear model (GLM) framework. The experiment had a slow event-related design, which was modeled using boxcar regressors convolved with a standard hemodynamic response function (HRF). An event was defined as a single vignette (20 s), and its onset was defined by the onset of the mental state segment of each vignette. The GLM also included six motion parameters as nuisance regressors.

Beta values were estimated in each voxel for all 3 conditions: cooperation, competition, and neutral. Contrast maps for the following contrasts were produced for each participant: (1) *cooperation and competition* > *neutral*, (2) *cooperation* > *competition*, (3) *competition* > *cooperation*. For each contrast, participants' images were used to perform group-level analyses. To correct for multiple comparisons, images from the group-level analyses were subjected to a voxel-wise threshold of $p < 0.001$ (uncorrected) and a cluster extent threshold ensuring $q < 0.05$ (FDR-corrected). Anatomical labels for peak coordinates were retrieved using SPM Anatomy Toolbox v2 (Eickhoff et al., 2005), and results were visualized using the xjView toolbox (http://www.alivelearn.net/xjview).

ROI analyses. The BOLD response over baseline to each condition was calculated for each ROI. Baseline response in each ROI was calculated as the average response in that ROI at all time points during the resting period, excluding the first 6 s after the offset of each stimulus (to allow the hemodynamic response to decay). The percent signal change (PSC) relative to baseline was calculated for each time point in each condition, averaging across all voxels in the ROI, where PSC (at time *t*) = 100 x [(average magnitude response for condition at time *t* – average magnitude response for fixation)/average magnitude response for fixation]. First, PSC was averaged across the mental state and question segments (5 TRs or 10 s; offset 6 s from presentation time to adjust for hemodynamic lag) to estimate a single PSC for each condition in each ROI for each participant. Linear mixed models were run using the package 'lme4' (Bates et al., 2015a) in R. We were primarily interested in understanding whether the PSC for ROIs differed across conditions. Our full model included condition (cooperation, competition, neutral) as a predictor. Participant and item were included as random effects, and we fit an intercept for each participant and for each item, allowing the intercept to vary across individuals and items. We performed a likelihood ratio tests (LRT) to test whether the model with our predictor of interest would provide a better fit to the data than a model without that term.

Models were also created with PSC as a predictor and behavioral measures (e.g., post-scan ratings) as response terms, with participant and item included as random effects.

Analyses of spatial patterns

Whole-brain analyses. Normalized (but unsmoothed) images were analyzed using a GLM framework. Again, onsets were defined by the onset of the mental state segment of each vignette, and six motion parameters were included as nuisance regressors. Betas were estimated in each voxel for each vignette. A whole-brain searchlight analysis was performed, with each searchlight containing 100 voxels. We used the CoSMoMVPA toolbox (Oosterhof, Connolly, & Haxby, 2016) to run the searchlight analysis. A Naïve bayes classifier was used with a 5-fold partitioning scheme for cross-validation. Participants' output searchlight images were used to perform group-level analyses. Group-level analyses were performed and corrected for multiple comparisons by computing a Monte Carlo based clustering statistic with Threshold-free cluster enhancement and 1000 permutations.

ROI analyses. A similar procedure to the searchlight procedure was used, except that instead of a searchlight sphere, ROIs were used. An accuracy score averaged across training/testing set combinations was computed for each ROI and every individual.

3.3 RESULTS

3.3.1 Behavioral analyses

Neither performance on the True/False questions nor reaction time differed across the three conditions (Figure 1; performance: $\chi^2(2) = 2.098$, $p = 0.35$; reaction time: $\chi^2(2) = 0.184$, $p = 0.990$). Post-scan ratings of niceness (Figure 2), however, differed across the three conditions such that ratings were higher for the cooperation condition than for the neutral condition (Wald test: $z = 12.14$, $p < 0.001$), and ratings for the neutral condition were higher than those for the competition condition (Wald test: $z = 14.77$, $p < 0.001$). These results indicate that our conditions—cooperation, competition, and neutral—were successfully designed: participants were able to detect differences between agents' cooperative, competitive, or neutral motivations.

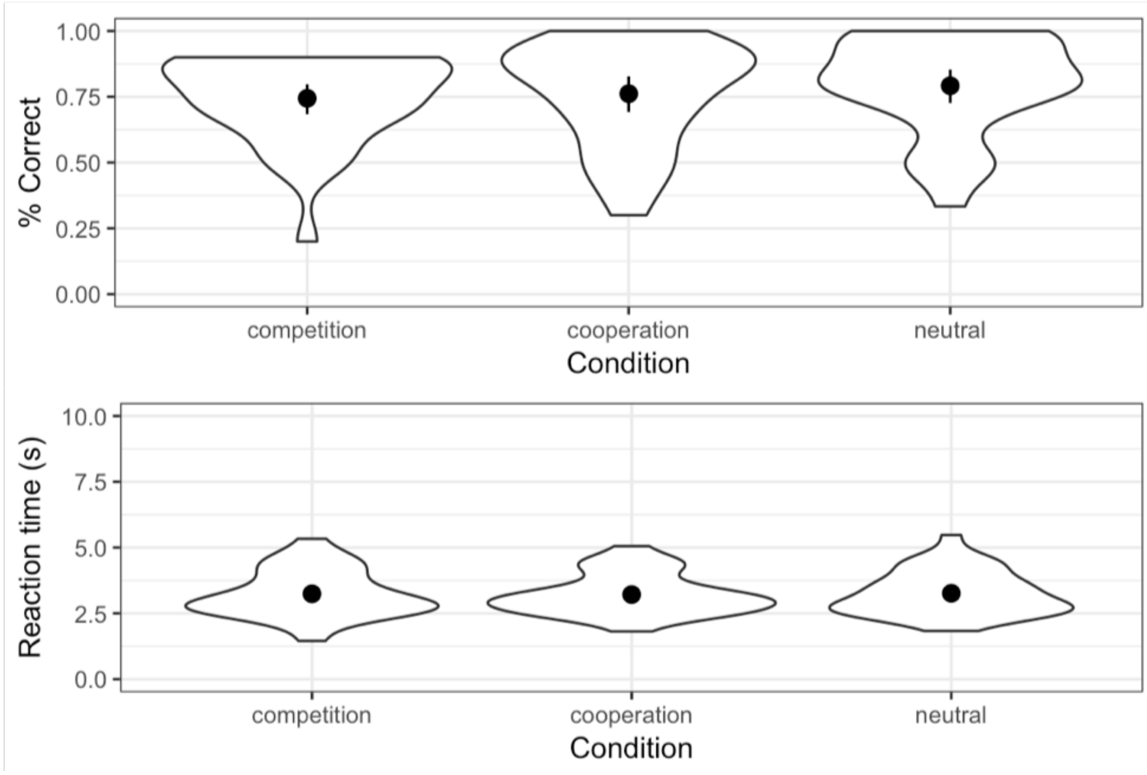


Figure 1. Performance (top) and reaction time (bottom) on the True/False question, split by condition. Violin plots display the distribution of performance and rating, respectively; error bars denote 95% CI.

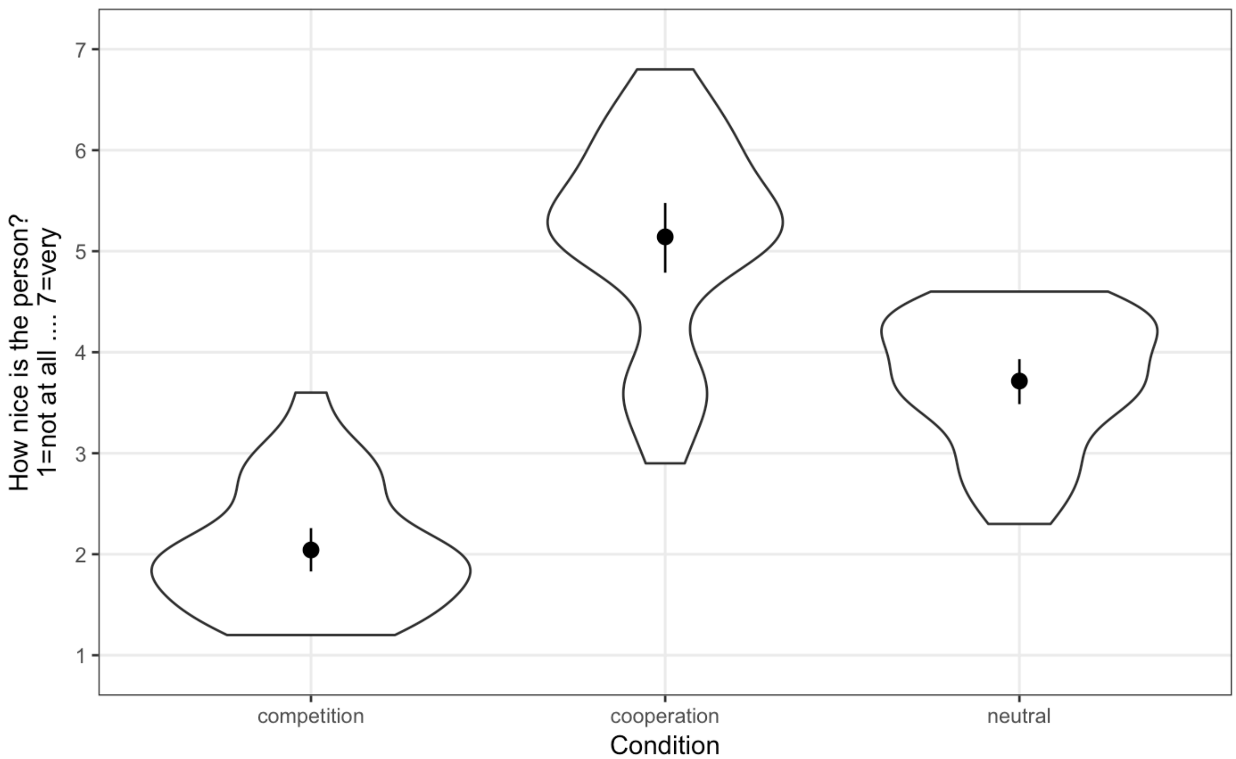


Figure 2. Post-scan ratings of how nice the agent is, split by condition. Violin plots display distribution of ratings; error bars denote 95% CI.

3.3.2 Analyses of response magnitudes

Whole-brain analyses

The contrasts *cooperation and competition* > *neutral* and *cooperation* > *competition* did not reveal any significant clusters. Meanwhile, the contrast *competition* > *cooperation* revealed two clusters, with peak coordinates in the left cerebellum ([-24, -70, -38]) and right medial temporal pole ([48 14 -38]).

ROI analyses

There was no significant difference in percent signal change across the three conditions when tested across ROIs; adding condition to the model did not significantly increase model fit ($\chi^2(2) = 1.242$, $p = 0.538$). Furthermore, there was no interaction between condition and ROI ($\chi^2(6) = 1.0005$, $p = 0.986$). We note that percent signal change here was calculated across the mental state and question segments as this window of time contained information relevant to the experimental condition. For the sake of completion, we also compared the effect of condition on percent signal change calculated across the entire trial as well as for each individual segment of the vignette—results reveal no main effect of condition and no interaction between condition and ROI (all *p*s > 0.05) (results for just the mental state segment is presented in Figure 3).

3.3.3 Analyses of spatial patterns

Whole-brain analyses

Searchlight analysis revealed one cluster for which the Naïve Bayes classifier significantly performed different from chance level (33%); this cluster had peak coordinates corresponding to the left superior orbital gyrus [-12, 44, -20; 31 voxels]. Similar results were achieved using a different classifier (i.e., LDA: [-12, 44, -20; 34 voxels]).

ROI analyses

The Naïve Bayes classifier was trained on all three types of vignettes; classification accuracy was no different from chance level (33%) for all four ROIs (one-sample *t*-test, *p*s > 0.05). We also replicated the same pattern of results with an LDA classifier, an SVM classifier, as well as an GNB classifier with time course data.

One possibility for this lack of finding could be due to the inclusion of the neutral vignettes. To test whether classification accuracy would be above chance level when faced with a binary classification problem, we trained the classifier using just cooperation and competition trials. Nonetheless, classification accuracies remained no different from chance level (50%), and this was the case even when we used different classifiers (*p*s > 0.05).

However, if we examine just the mental state segment, we get a different pattern of results regarding the comparison between cooperation and competition. We combined the ToM regions into one ROI and consistently found above-chance classification for this ROI with the LDA, SVM, and GNB classifiers (*p*s < 0.05). More importantly, this pattern was not consistently found for other comparisons (e.g., cooperation vs. competition vs. neutral; cooperation vs. neutral; competition vs. neutral).

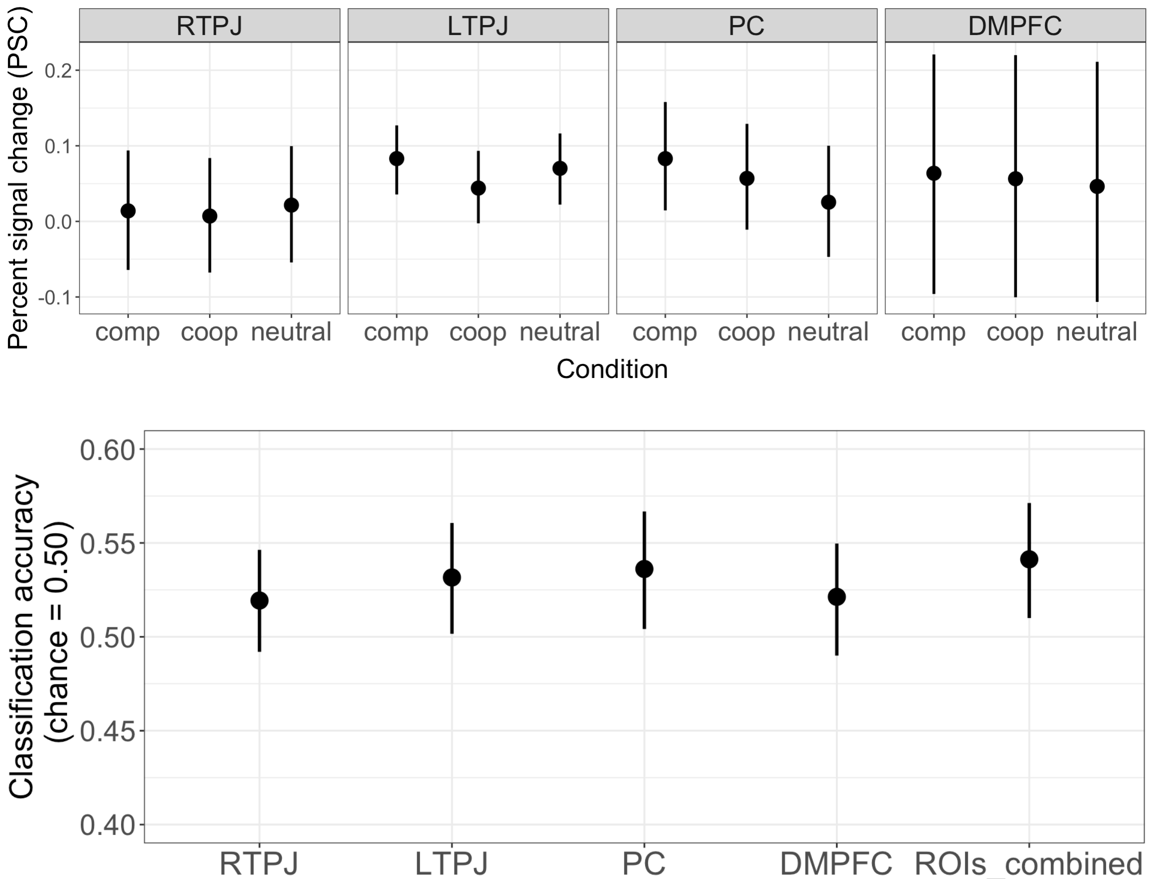


Figure 3. ROI results. Percent signal change across conditions (top) and classification accuracy for cooperation vs. competition (bottom) across ROIs for the mental state segment only. Error bars denote 95% CI.

3.3.4 Analyses on the relationship between neural data and behavioral data

Predicting post-scan ratings

First, we examined whether mean percent signal change (PSC) in any ROI predicted post-scan ratings of how nice the agent was in the vignette, and whether this differed across conditions (Figure 4). The three-way interaction between Condition, PSC, and ROI was significant ($\chi^2(6) = 45.555$, $p < 0.001$). Further comparisons revealed two-way interactions between PSC and Condition in RTPJ ($\chi^2(2) = 26.461$, $p < 0.001$), LTPJ ($\chi^2(2) = 13.695$, $p = 0.001$), and DMPFC ($\chi^2(2) = 10.846$, $p = 0.004$), but not PC ($\chi^2(2) = 2.207$, $p = 0.33$).

Once again, ratings of niceness were made on a scale of 1 (not at all) to 7 (very). In the RTPJ, an increase in PSC was associated with a decrease in the odds of making a rating of 1 (not at all) in the competition condition but not in any other condition. Similarly, an increase in PSC was associated with a decrease in the odds of making a rating of 7 (very) in the cooperation condition but not in any other condition. In the LTPJ, an increase in PSC was associated with an increase in the odds of making a rating of 1 (not at all) in the competition condition but not in any other condition. Likewise, an increase in PSC was associated with an increase in the odds of making a rating of 7 (very) in the cooperation condition but not in any other condition. In the DMPFC, an increase in PSC was associated with a increase in the odds of making a rating of 1 (not at all) in the competition condition but not in any other condition.

There were also significant three-way interactions when we examined PSC over just the mental state segment and question segment individually (mental state: $\chi^2(6) = 35.543$, $p < 0.001$; question: $\chi^2(6) = 26.678$, $p < 0.001$).

Second, we examined whether classification accuracy in any ROI predicted post-scan ratings of how nice the agent was in the vignette. Results suggest that classification accuracy in any ROI did not predict subsequent post-scan ratings ($ps > 0.05$).

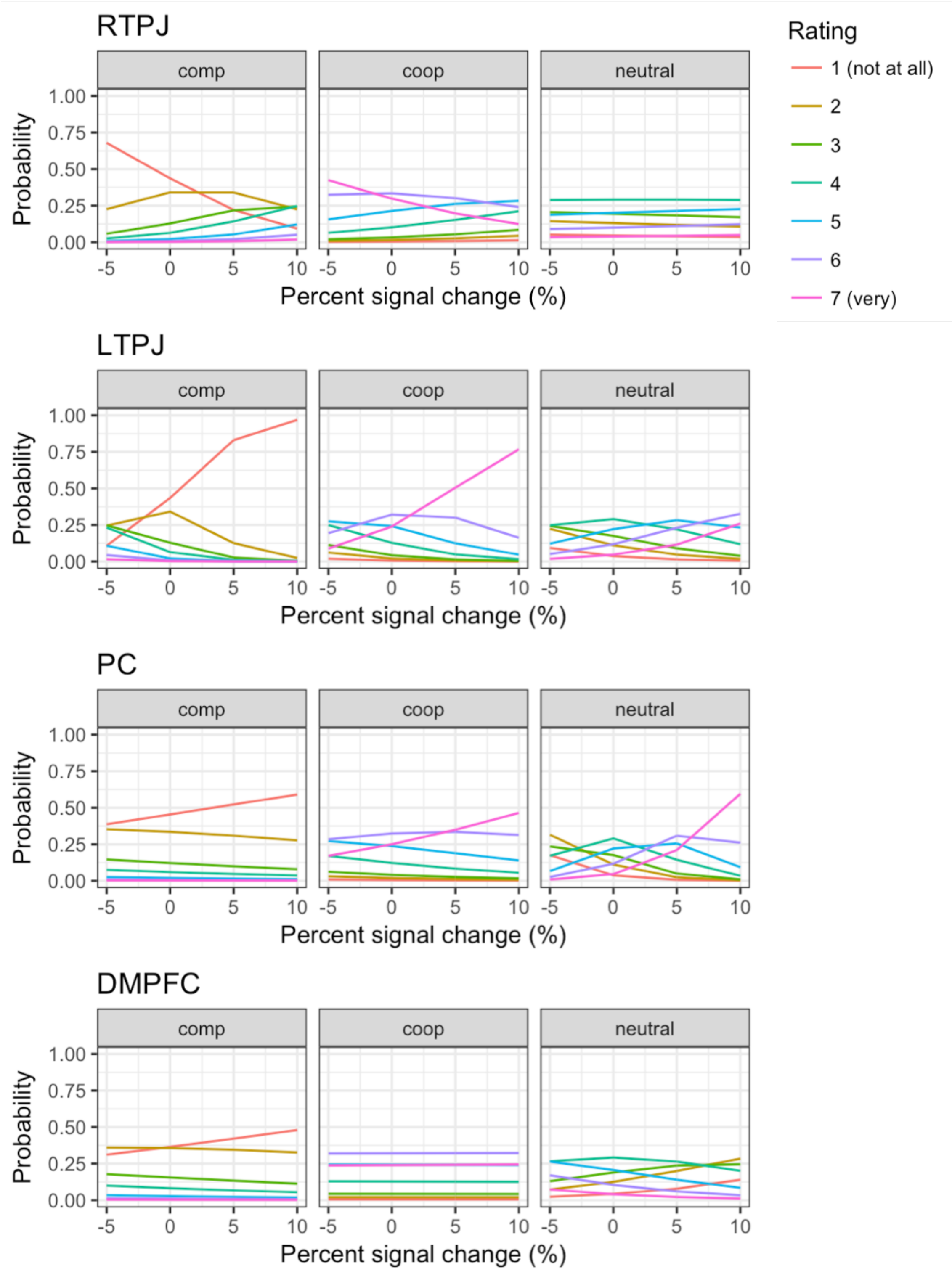


Figure 4. Relationship between PSC and probability of making a rating (as estimated by the model), separated by condition and ROI. Ratings were made from a scale of 1 (not at all) to 7 (very); rating was modeled as an ordered ordinal variable for all reported statistics.

Predicting behavioral performance on True/False questions

We also examined whether performance on the True/False questions was predicted by mean PSC for each ROI during the question segment of the vignette, and whether this differed across conditions and ROIs (Figure 5). Again, there was a three-way interaction between Condition, PSC, and ROI ($\chi^2(6) = 19.518$, $p = 0.003$). Two-way interactions between PSC and Condition were significant for the RTPJ ($\chi^2(2) = 6.165$, $p = 0.046$) and LTPJ ($\chi^2(2) = 13.627$, $p = 0.001$), but not for PC ($\chi^2(2) = 1.226$, $p = 0.542$) or DMPFC ($\chi^2(2) = 1.669$, $p = 0.434$).

Notably, greater activity in the RTPJ corresponded with better performance for the competition condition and worse performance in the cooperation and neutral conditions. Meanwhile, greater activity in the LTPJ corresponded with better performance for the competition and cooperation conditions and worse performance in the neutral condition.

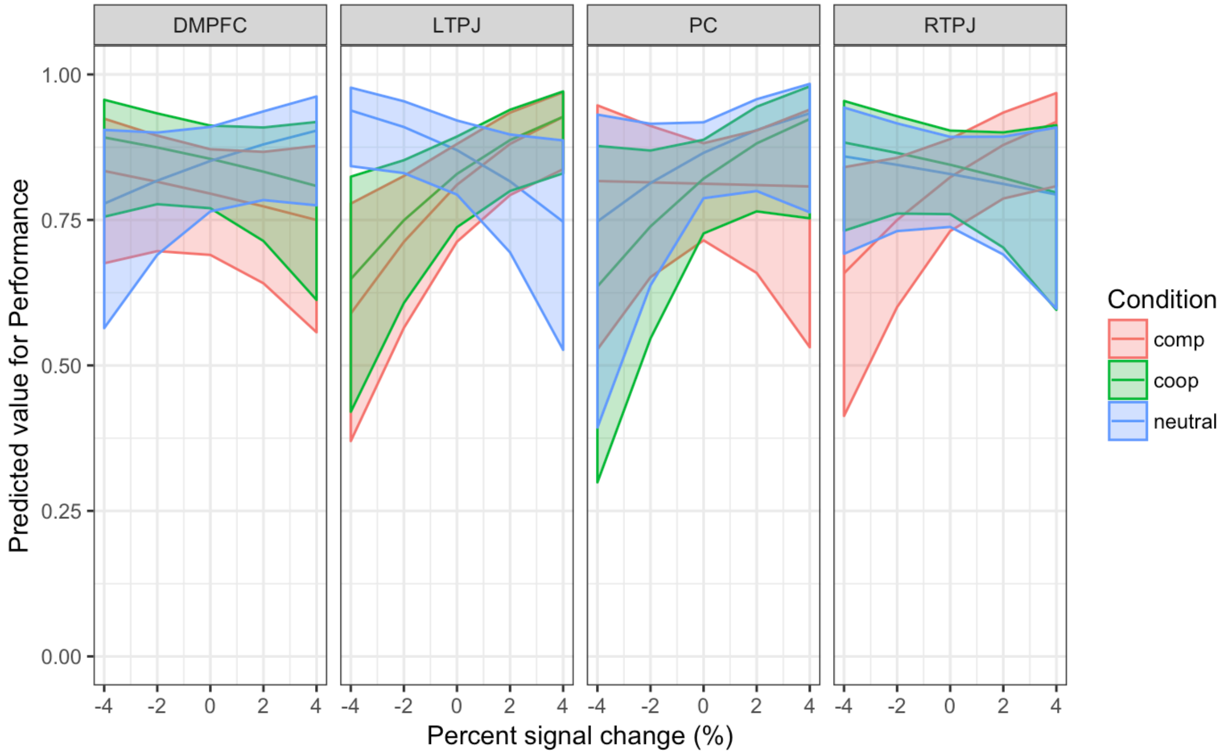


Figure 5. Relationship between predicted values for performance on True/False questions (as estimated by the model) and PSC, separated by ROI and Condition. Ribbons denote 95% CI.

We also examined whether classification accuracy in any ROI predicted overall performance on the True/False questions; results suggest that classification accuracy in any ROI did not predict overall performance (*p*s > 0.05).

3.4 DISCUSSION

In this study, we examined whether people engaged in ToM to a greater extent for competitive versus cooperative contexts or vice versa. Participants in the MRI scanner read vignettes about agents that acted in a cooperative, competitive, or neutral manner toward another person. Overall response magnitudes in regions implicated in ToM revealed no difference in response magnitude across conditions, suggesting that people do not engage in ToM more for competition than cooperation or vice versa. More importantly, these regions encode information separating cooperation and competition at the time during which this information is presented, even though such information is irrelevant for the actual task.

The univariate result conflicts with prior work that found greater medial prefrontal cortex activity for competition or deception when compared to cooperation (Decety et al., 2004; Lissek et al., 2008). However, the competition/deception conditions in these studies may have made greater mentalizing demands—that is, people had to check for mismatches in their own expectations and others' intentions in order to compete or deceive, whereas those types of checks were not needed in the cooperation conditions. This type of concern did not apply to the current study—in all conditions, participants had to reason about the false beliefs of the agent in the vignette. The results of the present study instead converge with prior work in our laboratory revealing no difference in response magnitude across cooperative and competitive interactions with the same partner; this study, too, was tightly controlled and was not susceptible to the same type of concern facing the studies mentioned above (Tsoi et al., 2016). We also do not think this null finding is a result of the study being underpowered as we had performed a power analysis to determine our sample size prior to data collection. Instead, we suggest that adults do not engage in ToM to different extents across cooperation and competition when all else are equal.

Intriguingly, ToM regions encoded information separating cooperation and competition during the timepoints in which such information was presented, which converge with the results of a prior study revealing sensitivity to differences between cooperation and competition in an interactive game (Tsoi et al., 2016). In this prior study, information about context was integral to the task at hand—participants had to achieve a cooperative or competitive goal. In contrast, participants in the current study were passive observers of others' cooperative and competitive behaviors; furthermore, participants did not rely at all on information about context to do well in the experimental task. Another difference between the two studies is that people in the prior study played the game with just one other person, thereby cooperating and competing with the same person, whereas in the present study, people were only presented with one behavior for each agent and were never exposed to different behaviors by the same agent. Despite many differences between the two tasks, the results are remarkably similar. In the domain of moral judgments, information related to the intentional/accidental dimension of actions (Koster-Hale, Saxe, Dungan, & Young, 2013) and physical/psychological dimension of harmful actions (Tsoi, Dungan, Chakroff, & Young, 2018) are encoded in regions implicated in ToM. In the domain of social interaction, information about context may be encoded regardless of whether the information is relevant to the self.

We also examined the relationship between neural data and post-scan behavioral ratings of agents in the vignettes. Intriguingly, we found a three-way interaction between

Condition, PSC, and ROI. Specifically, we found that increases in activation in the RTPJ led to lower odds of making an extreme rating—that is, greater RTPJ activity led to a decreased likelihood of saying that the agent in the competition condition was not at all nice and a decreased likelihood of saying that the agent in the cooperation condition was very nice. We observed the opposite relationship in the LTPJ: greater LTPJ activity led to a greater likelihood of saying that the agent in the competition was not at all nice and that the agent in the cooperation condition was very nice. The DMPFC only showed an effect in the competition condition, in which greater activity in the DMPC led to an increased likelihood of saying that the agent in the competition condition was not at all nice. The opposite patterns found for the RTPJ and LTPJ support the idea that the function of the two regions in social cognition may be distinct. While the idea is not novel, the literature on this has not provided a consistent story. That is, while some studies report bilateral TPJ activation (Gallagher et al., 2000; Jenkins & Mitchell, 2010; Santiesteban, Banissy, Catmur, & Bird, 2015), others have revealed a selective role of RTPJ in belief attributions in the domain of moral judgment (Aichhorn et al., 2006; Koster-Hale et al., 2013; Saxe & Wexler, 2005; Saxe & Powell, 2006; Young et al., 2007) and a selective role of LTPJ in visual perspective-taking (Schurz, Aichhorn, Martin, & Perner, 2013).

One interpretation for the difference we found across RTPJ and LTPJ is that the LTPJ may be tracking with perceptions of how people may initially think about an agent, and the RTPJ may be carrying out computations of whether one should temper their judgments of others at either extreme end. However, this interpretation goes against prior work showing that RTPJ activity positively predicts praise ratings for good vs. bad actions and blame ratings for bad vs. good actions (Yoder & Decety, 2014), as well as other work showing that people with higher RTPJ response to accidental harm are more forgiving and give out less blame as compared to people with lower RTPJ response to accidental harm (Young & Saxe, 2009b). More investigations of how neural data relates to social behavior will likely be fruitful in characterizing the different roles for the individual regions that comprise the ToM network.

Lastly, we also examined whether activity in the ToM network predicted whether someone would respond to the True/False question correctly, and whether that relationship differed across condition. The RTPJ and LTPJ showed somewhat similar patterns of results: activity in these regions tracked negatively with performance in the neutral condition, but they positively tracked with performance in the competition condition. Where the two regions differed was in their relationship with performance in the cooperation condition: the relationship between LTPJ activity and performance in the cooperation condition was more similar to the relationship between LTPJ activity and performance in the competition condition, whereas the relationship between RTPJ activity and performance in the cooperation condition was more similar to the relationship between RTPJ activity and performance in the neutral condition. The negative relationships were puzzling and goes against the idea that greater ToM activity would result in greater success in reasoning about people's mental states. Even more puzzling is the finding that the direction of the relationship differed across conditions. One potential explanation for this result is that people may be overattributing intentions (good or bad) to agents in the neutral condition, which would distract them from answering the question correctly. This explanation would also predict longer reaction times in the neutral condition, but reaction times did not differ across conditions. More work is needed to question the idea that greater activity in ToM regions reflects greater accuracy in mental state understanding—or how accurate mental state reasoning can be supported by a two-system account of ToM in which a more quick, intuitive system is pitted against a slower, deliberate system (Butterfill & Apperly, 2013; but see Carruthers, 2016), similar to what has been proposed to support moral reasoning (Greene et al., 2001; Greene, 2013) and general decision making (Kahneman, 2012).

Together, results from this study support the idea that while overall levels of activity within regions implicated in ToM do not differ across social contexts, activity in these regions track with accurate belief understanding and subsequent impressions of others and vary depending on the context.

3.5 REFERENCES

Agresti, A., & Coull, B. A. (1998). Approximate is better than "exact" for interval estimation of binomial proportions. *The American Statistician*, *52*(2), 119–126.

Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Do visual perspective tasks need theory of mind? *NeuroImage*, *30*(3), 1059–1068. https://doi.org/10.1016/j.neuroimage.2005.10.026

Aloise-Young, P. A. (1993). The development of self-presentation: Self-promotion in 6- to 10-year-old children. *Social Cognition*, *11*(2), 201–222. https://doi.org/10.1521/soco.1993.11.2.201

Andreoni, J., & Bernheim, B. D. (2009). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, *77*(5), 1607–1636. https://doi.org/10.3982/ECTA7384

Apfelbaum, E. P., Pauker, K., Ambady, N., Sommers, S. R., & Norton, M. I. (2008). Learning (not) to talk about race: When older children underperform in social categorization. *Developmental Psychology*, *44*(5), 1513–1518. https://doi.org/10.1037/a0012835

Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: A 1-year longitudinal perspective. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00655

Banerjee, R. (2002). Audience effects on self-presentation in childhood. *Social Development*, *11*(4), 487–507. https://doi.org/10.1111/1467-9507.00212

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*(2), 241–251.

Baron-Cohen, Simon. (2000). Theory of mind and autism: A fifteen year review. *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience (2nd Ed.)*, 3–20.

Baron-Cohen, Simon, Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*(1), 37–46.

Baron-Cohen, Simon, Wheelwright, S., Spong, A., Scahill, V., & Lawson, J. (2001). Are intuitive physics and intuitive psychology independent? A test with children with Asperger Syndrome. *Journal of Developmental and Learning Disorders*, *5*(1), 47–78.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015a). Fitting linear mixed-effects models using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015b). Fitting linear mixed-effects models using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Batson, C. D., Thompson, E. R., & Chen, H. (2002). Moral hypocrisy: Addressing some alternatives. *Journal of Personality and Social Psychology*, *83*(2), 330–339. https://doi.org/10.1037/0022-3514.83.2.330

Baumard, N., Mascaro, O., & Chevallier, C. (2012). Preschoolers are able to take merit into account when distributing goods. *Developmental Psychology*, *48*(2), 492–498. https://doi.org/10.1037/a0026598

Begeer, S., Rieffe, C., Terwogt, M. M., & Stockmann, L. (2003). Theory of mind–based action in children from the autism spectrum. *Journal of Autism and Developmental Disorders*, *33*(5), 479–487.

Blake, P. R., McAuliffe, K., Corbit, J., Callaghan, T. C., Barry, O., Bowie, A., … Warneken, F. (2015). The ontogeny of fairness in seven societies. *Nature*, *528*(7581), 258–261. https://doi.org/10.1038/nature15703

Blake, Peter R., & McAuliffe, K. (2011). "I had so much it didn't seem fair": Eight-year-olds reject two forms of inequity. *Cognition*, *120*(2), 215–224. https://doi.org/10.1016/j.cognition.2011.04.006

Blake, Peter R., McAuliffe, K., & Warneken, F. (2014). The developmental origins of fairness: the knowledge–behavior gap. *Trends in Cognitive Sciences*, *18*(11), 559–561. https://doi.org/10.1016/j.tics.2014.08.003

Blake, Peter R., Piovesan, M., Montinari, N., Warneken, F., & Gino, F. (2015). Prosocial norms in the classroom: The role of self-regulation in following norms of giving. *Journal of Economic Behavior & Organization*, *115*, 18–29. https://doi.org/10.1016/j.jebo.2014.10.004

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, *77*(1), B25-31.

Brosnan, S. F., & Bshary, R. (2016). On potential links between inequity aversion and the structure of interactions for the evolution of cooperation. *Behaviour*, *153*(9–11), 1267–1292. https://doi.org/10.1163/1568539X-00003355

Bruneau, E. G., & Saxe, R. (2012). The power of being heard: The benefits of 'perspective-giving' in the context of intergroup conflict. *Journal of Experimental Social Psychology*, *48*(4), 855–866. https://doi.org/10.1016/j.jesp.2012.02.017

Buhrmester, D., Goldfarb, J., & Cantrell, D. (1992). Self-presentation when sharing with friends and nonfriends. *The Journal of Early Adolescence*, *12*(1), 61–79.

Bull, R., Phillips, L. H., & Conway, C. A. (2008). The role of control functions in mentalizing: Dual-task studies of Theory of Mind and executive function.

Cognition, *107*(2), 663–672. https://doi.org/10.1016/j.cognition.2007.07.015

Burkart, J. M., Hrdy, S. B., & Van Schaik, C. P. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, *18*(5), 175–186. https://doi.org/10.1002/evan.20222

Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language*, *28*(5), 606–637. https://doi.org/10.1111/mila.12036

Byrne, R. W., & Corp, N. (2004). Neocortex size predicts deception rate in primates. *Proceedings of the Royal Society B: Biological Sciences*, *271*(1549), 1693–1699. https://doi.org/10.1098/rspb.2004.2780

Byrne, Richard W., & Whiten, A. (Eds.). (1988). *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford : New York: Clarendon Press ; Oxford University Press.

Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., & Eickhoff, S. B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure and Function*, *217*(4), 783–796. https://doi.org/10.1007/s00429-012-0380-y

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*(4), 1032–1053.

Carlson, Stephanie M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development*, *11*(2), 73–92. https://doi.org/10.1002/icd.298

Carlson, Stephanie M, Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, *87*(4), 299–319. https://doi.org/10.1016/j.jecp.2004.01.002

Carruthers, P. (2016). Two Systems for Mindreading? *Review of Philosophy and Psychology*, *7*(1), 141–162. https://doi.org/10.1007/s13164-015-0259-y

Chandler, M., Fritz, A. S., & Hala, S. (1989). Small-Scale Deceit: Deception as a Marker of Two-, Three-, and Four-Year-Olds' Early Theories of Mind. *Child Development*, *60*(6), 1263. https://doi.org/10.2307/1130919

Christensen, R. H. B. (2015). *ordinal—Regression Models for Ordinal Data*.

Cikara, M., Bruneau, E. G., & Saxe, R. R. (2011). Us and Them: Intergroup Failures of Empathy. *Current Directions in Psychological Science*, *20*(3), 149–153. https://doi.org/10.1177/0963721411408713

Corbit, J., McAuliffe, K., Callaghan, T. C., Blake, P. R., & Warneken, F. (2017). Children's collaboration induces fairness rather than generosity. *Cognition*, *168*, 344–356. https://doi.org/10.1016/j.cognition.2017.07.006

Cortes, B. P., Demoulin, S., Rodriguez, R. T., Rodriguez, A. P., & Leyens, J.-P. (2005). Infrahumanization or Familiarity? Attribution of Uniquely Human Emotions to the Self, the Ingroup, and the Outgroup. *Personality and Social Psychology Bulletin*, *31*(2), 243–253. https://doi.org/10.1177/0146167204271421

Dadds, M. R., Hunter, K., Hawes, D. J., Frost, A. D. J., Vassallo, S., Bunn, P., … Masry, Y. E. (2008). A measure of cognitive and affective empathy in children using parent ratings. *Child Psychiatry and Human Development, 39*(2), 111–122. https://doi.org/10.1007/s10578-007-0075-4

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, *33*(1), 67–80. https://doi.org/10.1007/s00199-006-0153-z

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. *Nature*, *446*(7137), 794–796. https://doi.org/10.1038/nature05651

Decety, J., & Cacioppo, S. (2012). The speed of morality: A high-density electrical neuroimaging study. *Journal of Neurophysiology*, *108*(11), 3068–3072. https://doi.org/10.1152/jn.00473.2012

Decety, Jean, Jackson, P. L., Sommerville, J. A., Chaminade, T., & Meltzoff, A. N. (2004). The neural bases of cooperation and competition: An fMRI investigation. *NeuroImage*, *23*(2), 744–751. https://doi.org/10.1016/j.neuroimage.2004.05.025

Devine, R. T., & Hughes, C. (2016). Measuring theory of mind across middle childhood: Reliability and validity of the Silent Films and Strange Stories tasks. *Journal of Experimental Child Psychology*, *149*, 23–40. https://doi.org/10.1016/j.jecp.2015.07.011

Devine, R. T., White, N., Ensor, R., & Hughes, C. (2016). Theory of mind in middle childhood: Longitudinal associations with executive function and social competence. *Developmental Psychology*, *52*(5), 758–771. https://doi.org/10.1037/dev0000105

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *NeuroImage*, *55*(2), 705–712. https://doi.org/10.1016/j.neuroimage.2010.12.040

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, *25*(4), 1325–1335. https://doi.org/10.1016/j.neuroimage.2004.12.034

Engelmann, J. M., Herrmann, E., & Tomasello, M. (2012). Five-year olds, but not chimpanzees, attempt to manage their reputations. *PLoS ONE*, *7*(10), e48433. https://doi.org/10.1371/journal.pone.0048433

Engelmann, J. M., & Rapp, D. J. (2018). The influence of reputational concerns on children's prosociality. *Current Opinion in Psychology*, *20*, 92–95. https://doi.org/10.1016/j.copsyc.2017.08.024

Epley, N., Akalis, S., Waytz, A., & Cacioppo, J. T. (2008). Creating Social Connection Through Inferential Reproduction: Loneliness and Perceived Agency in Gadgets, Gods, and Greyhounds. *Psychological Science*, *19*(2), 114–120. https://doi.org/10.1111/j.1467-9280.2008.02056.x

Epley, N., & Waytz, A. (2010). Mind Perception. In *Handbook of Social Psychology*. John Wiley & Sons, Inc. Retrieved from http://dx.doi.org/10.1002/9780470561119.socpsy001014

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868.

FeldmanHall, O., Dalgleish, T., Thompson, R., Evans, D., Schweizer, S., & Mobbs, D. (2012). Differential neural circuitry and self-interest in real *vs* hypothetical moral decisions. *Social Cognitive and Affective Neuroscience*, *7*(7), 743–751. https://doi.org/10.1093/scan/nss069

FeldmanHall, O., Mobbs, D., Evans, D., Hiscox, L., Navrady, L., & Dalgleish, T. (2012). What we say and what we do: The relationship between real and hypothetical moral choices. *Cognition*, *123*(3), 434–441. https://doi.org/10.1016/j.cognition.2012.02.001

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly*, *50*(3), 274–290.

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition*, *57*(2), 109–128.

Flombaum, J. I., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology*, *15*(5), 447–452. https://doi.org/10.1016/j.cub.2004.12.076

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*(4), 708–724.

Galinsky, Adam D., Maddux, W. W., Gilin, D., & White, J. B. (2008). Why it pays to get inside the head of your opponent: The differential effects of perspective taking and empathy in negotiations. *Psychological Science*, *19*(4), 378–384. https://doi.org/10.1111/j.1467-9280.2008.02096.x

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of "theory of mind" in verbal and nonverbal tasks. *Neuropsychologia*, *38*(1), 11–21.

Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental Science*, *14*(5), 1012–1020. https://doi.org/10.1111/j.1467-7687.2011.01048.x

Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J., & Haxby, J. V. (2007). Two takes on the social brain: A comparison of theory of mind tasks. *Journal of Cognitive Neuroscience*, *19*(11), 1803–1814. https://doi.org/10.1162/jocn.2007.19.11.1803

Gonzalez, Antonya M., Steele, J. R., & Baron, A. S. (2017). Reducing children's implicit racial bias through exposure to positive out-group exemplars. *Child Development*, *88*(1), 123–130. https://doi.org/10.1111/cdev.12582

Gonzalez, Antonya Marie, Dunlop, W. L., & Baron, A. S. (2016). Malleability of implicit associations across development. *Developmental Science*. https://doi.org/10.1111/desc.12481

Greene, Joshua D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105–2108.

Greene, Joshua D. (2013). *Moral tribes: Emotion, reason, and the gap between us and them*. New York: The Penguin Press.

Hadwin, J., & Perner, J. (1991). Pleased and surprised: Children's cognitive theory of emotion. *British Journal of Developmental Psychology*, *9*(2), 215–234. https://doi.org/10.1111/j.2044-835X.1991.tb00872.x

Hala, S., Chandler, M., & Fritz, A. S. (1991). Fledgling theories of mind: Deception as a marker of three-year-olds' understanding of false belief. *Child Development*, *62*(1), 83–97. https://doi.org/10.1111/j.1467-8624.1991.tb01516.x

Hamann, K., Bender, J., & Tomasello, M. (2014). Meritocratic sharing is based on collaboration in 3-year-olds. *Developmental Psychology*, *50*(1), 121–128. https://doi.org/10.1037/a0032965

Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature*, *476*(7360), 328–331. https://doi.org/10.1038/nature10278

Happé, F. G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, *66*(3), 843–855.

Hare, B. (2001). Can competitive paradigms increase the validity of experiments on primate social cognition? *Animal Cognition*, *4*(3–4), 269–280. https://doi.org/10.1007/s100710100084

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*(4), 771–785. https://doi.org/10.1006/anbe.1999.1377

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, *61*(1), 139–151. https://doi.org/10.1006/anbe.2000.1518

Hare, B., Call, J., & Tomasello, M. (2006). Chimpanzees deceive a human competitor by hiding. *Cognition*, *101*(3), 495–514. https://doi.org/10.1016/j.cognition.2005.01.011

Hare, B., & Tomasello, M. (2004). Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. *Animal Behaviour*, *68*(3), 571–581. https://doi.org/10.1016/j.anbehav.2003.11.011

Harris, L. T., & Fiske, S. T. (2007). Social groups that elicit disgust are differentially processed in mPFC. *Social Cognitive and Affective Neuroscience*, *2*(1), 45–51. https://doi.org/10.1093/scan/nsl037

Harris, L. T., & Fiske, S. T. (2009). Social neuroscience evidence for dehumanised perception. *European Review of Social Psychology*, *20*(1), 192–231. https://doi.org/10.1080/10463280902954988

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, *10*(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4

Haslam, N., & Loughnan, S. (2014). Dehumanization and infrahumanization. *Annual Review of Psychology*, *65*(1), 399–423. https://doi.org/10.1146/annurev-psych-010213-115045

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, *57*(2), 243. https://doi.org/10.2307/1416950

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Heyes, C. M. (1998). Theory of mind in nonhuman primates. *The Behavioral and Brain Sciences*, *21*(1), 101–114; discussion 115-148.

Hrdy, S. B. (2009). *Mothers and others: The evolutionary origins of mutual understanding*. Cambridge, Mass: Belknap Press of Harvard University Press.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, *16*(2), 233–253. https://doi.org/10.1111/j.2044-835X.1998.tb00921.x

Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development*, *13*(4), 590–619.

Jackson, J. W. (1993). Realistic group conflict theory: A review and evaluation of the theoretical and empirical literature. *The Psychological Record*, *43*(3), 395–413.

Jenkins, A. C., & Mitchell, J. P. (2010). Mentalizing under uncertainty: Dissociated neural responses to ambiguous and unambiguous mental state inferences. *Cerebral Cortex*, *20*(2), 404–410. https://doi.org/10.1093/cercor/bhp109

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Kahneman, D. (2012). *Thinking, fast and slow*. London: Penguin Books.

Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition*, *109*(2), 224–234. https://doi.org/10.1016/j.cognition.2008.08.010

Kanngiesser, P., & Warneken, F. (2012). Young children consider merit when sharing resources with others. *PLoS ONE*, *7*(8), e43979. https://doi.org/10.1371/journal.pone.0043979

Karg, K., Schmelz, M., Call, J., & Tomasello, M. (2015). The goggles experiment: can chimpanzees use self-experience to infer what a competitor can see? *Animal Behaviour*, *105*, 211–221. https://doi.org/10.1016/j.anbehav.2015.04.028

Kelman, H. G. (1973). Violence without moral restraint: Reflections on the dehumanization of victims and victimizers. *Journal of Social Issues*, *29*(4), 25–61. https://doi.org/10.1111/j.1540-4560.1973.tb00102.x

Koster-Hale, J., Saxe, R., Dungan, J., & Young, L. L. (2013). Decoding moral judgments from neural representations of intentions. *Proceedings of the National Academy of Sciences*, *110*(14), 5648–5653. https://doi.org/10.1073/pnas.1207992110

Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology*, *90*(4), 543–555. https://doi.org/10.1037/0022-3514.90.4.543

Lee, K. (2013). Little Liars: Development of Verbal Deception in Children. *Child Development Perspectives*, *7*(2), 91–96. https://doi.org/10.1111/cdep.12023

Leimgruber, K. L., Shaw, A., Santos, L. R., & Olson, K. R. (2012). Young children are more generous when others are aware of their actions. *PLoS ONE*, *7*(10), e48292. https://doi.org/10.1371/journal.pone.0048292

Leyens, J.-P., Paladino, P. M., Rodriguez-Torres, R., Vaes, J., Demoulin, S., Rodriguez-Perez, A., & Gaunt, R. (2000). The emotional side of prejudice: The attribution of secondary emotions to ingroups and outgroups. *Personality and Social Psychology Review*, *4*(2), 186–197. https://doi.org/10.1207/S15327957PSPR0402_06

Lissek, S., Peters, S., Fuchs, N., Witthaus, H., Nicolas, V., Tegenthoff, M., … Brüne, M. (2008). Cooperation and deception recruit different subsets of the theory-of-mind network. *PLoS ONE*, *3*(4), e2023. https://doi.org/10.1371/journal.pone.0002023

Liu, B., Huang, Z., Xu, G., Jin, Y., Chen, Y., Li, X., … Jing, J. (2016). Altruistic sharing behavior in children: Role of theory of mind and inhibitory control. *Journal of Experimental Child Psychology*, *141*, 222–228. https://doi.org/10.1016/j.jecp.2015.09.010

Loewenstein, G. F., Thompson, L., & Bazerman, M. H. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social Psychology*, *57*(3), 426.

Lurz, R. W. (2011). *Mindreading animals: The debate over what animals know about other minds*. Cambridge, Mass: MIT Press.

Lyons, D. E., & Santos, L. R. (2006). Ecology, domain specificity, and the origins of theory of mind: Is competition the catalyst? *Philosophy Compass*, *1*(5), 481–492. https://doi.org/10.1111/j.1747-9991.2006.00032.x

MacLaren, R., & Olson, D. (1993). Trick or treat: Children's understanding of surprise. *Cognitive Development*, *8*(1), 27–46. https://doi.org/10.1016/0885-2014(93)90003-N

Malle, B. F., & Holbrook, J. (2012). Is there a hierarchy of social inferences? The likelihood and speed of inferring intentionality, mind, and personality. *Journal of Personality and Social Psychology*, *102*(4), 661–684. https://doi.org/10.1037/a0026790

McAuliffe, K., Blake, P. R., Kim, G., Wrangham, R. W., & Warneken, F. (2013). Social influences on inequity aversion in children. *PLoS ONE*, *8*(12), e80966. https://doi.org/10.1371/journal.pone.0080966

McAuliffe, K., Blake, P. R., Steinbeis, N., & Warneken, F. (2017). The developmental foundations of human fairness. *Nature Human Behaviour*, *1*, 0042.

McAuliffe, K., & Dunham, Y. (2016). Group bias in cooperative norm enforcement. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150073. https://doi.org/10.1098/rstb.2015.0073

McPherson Frantz, C., & Janoff-Bulman, R. (2000). Considering Both Sides: The Limits of Perspective Taking. *Basic and Applied Social Psychology*, *22*(1), 31–42. https://doi.org/10.1207/S15324834BASP2201_4

Melis, A. P., Call, J., & Tomasello, M. (2006). Chimpanzees (Pan troglodytes) conceal visual and auditory information from others. *Journal of Comparative Psychology*, *120*(2), 154–162. https://doi.org/10.1037/0735-7036.120.2.154

Morewedge, C. K. (2009). Negativity bias in attribution of external agency. *Journal of Experimental Psychology: General*, *138*(4), 535–545. https://doi.org/10.1037/a0016796

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science (New York, N.Y.)*, *308*(5719), 255–258. https://doi.org/10.1126/science.1107621

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. *Frontiers in Neuroinformatics*, *10*. https://doi.org/10.3389/fninf.2016.00027

Opotow, S. (1990). Moral exclusion and injustice: An introduction. *Journal of Social Issues*, *46*(1), 1–20. https://doi.org/10.1111/j.1540-4560.1990.tb00268.x

Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology*, *1*(1), 33–60.

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental Psychology*, *43*(4), 974–990. https://doi.org/10.1037/0012-1649.43.4.974

Perner, J, & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences*, *3*(9), 337–344.

Perner, Josef, & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences*, *16*(10), 519–525. https://doi.org/10.1016/j.tics.2012.08.004

Peterson, C. C., Slaughter, V., Peterson, J., & Premack, D. (2013). Children with autism can track others' beliefs in a competitive game. *Developmental Science*, *16*(3), 443–450. https://doi.org/10.1111/desc.12040

Phillips, W., Barnes, J. L., Mahajan, N., Yamaguchi, M., & Santos, L. R. (2009). 'Unwilling' versus 'unable': capuchin monkeys' (*Cebus apella*) understanding of human intentional action: Unwilling vs. unable in capuchin monkeys. *Developmental Science*, *12*(6), 938–945. https://doi.org/10.1111/j.1467-7687.2009.00840.x

Povinelli, D. J., Rulf, A. B., & Bierschwale, D. T. (1994). Absence of knowledge attribution and self-recognition in young chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *108*(1), 74–80. https://doi.org/10.1037/0735-7036.108.1.74

Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: suspiciously human? *Trends in Cognitive Sciences*, *7*(4), 157–160. https://doi.org/10.1016/S1364-6613(03)00053-6

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(04), 515. https://doi.org/10.1017/S0140525X00076512

Raihani, N. J., & McAuliffe, K. (2012). Human punishment is motivated by inequity aversion, not a desire for reciprocity. *Biology Letters*, *8*(5), 802–804. https://doi.org/10.1098/rsbl.2012.0470

Reis, H. T., & Gruzen, J. (1976). On mediating equity, equality, and self-interest: The role of self-presentation in social exchange. *Journal of Experimental Social Psychology*, *12*(5), 487–503. https://doi.org/10.1016/0022-1031(76)90079-2

Rizzo, M. T., & Killen, M. (2016). Children's understanding of equity in the context of inequality. *British Journal of Developmental Psychology*, *34*(4), 569–581. https://doi.org/10.1111/bjdp.12150

Rochat, P., Dias, M. D. G., Guo Liping, Broesch, T., Passos-Ferreira, C., Winning, A., & Berg, B. (2009). Fairness in distributive justice by 3- and 5-year-olds across seven cultures. *Journal of Cross-Cultural Psychology*, *40*(3), 416–442. https://doi.org/10.1177/0022022109332844

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and U.S. preschoolers. *Psychological Science*, *17*(1), 74–81. https://doi.org/10.1111/j.1467-9280.2005.01667.x

Santiesteban, I., Banissy, M. J., Catmur, C., & Bird, G. (2015). Functional lateralization of temporoparietal junction - imitation inhibition, visual perspective-taking and theory of mind. *European Journal of Neuroscience*, *42*(8), 2527–2533. https://doi.org/10.1111/ejn.13036

Santos, L. R., Nissen, A. G., & Ferrugia, J. A. (2006). Rhesus monkeys, Macaca mulatta, know what others can and cannot hear. *Animal Behaviour*, *71*(5), 1175–1181. https://doi.org/10.1016/j.anbehav.2005.10.007

Saxe, R, & Kanwisher, N. (2003). People thinking about thinking peopleThe role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*(4), 1835–1842. https://doi.org/10.1016/S1053-8119(03)00230-1

Saxe, R, & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, *43*(10), 1391–1399. https://doi.org/10.1016/j.neuropsychologia.2005.02.013

Saxe, Rebecca, Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, *55*, 87–124. https://doi.org/10.1146/annurev.psych.55.090902.142044

Saxe, Rebecca, & Powell, L. J. (2006). It's the thought that counts: Specific brain regions for one component of theory of mind. *Psychological Science*, *17*(8), 692–699. https://doi.org/10.1111/j.1467-9280.2006.01768.x

Schmelz, M., Call, J., & Tomasello, M. (2011). Chimpanzees know that others make inferences. *Proceedings of the National Academy of Sciences*, *108*(7), 3077–3079. https://doi.org/10.1073/pnas.1000469108

Schmelz, Martin, & Call, J. (2016). The psychology of primate cooperation and competition: a call for realigning research agendas. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150067. https://doi.org/10.1098/rstb.2015.0067

Schmidt, M. F. H., & Sommerville, J. A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. *PLoS ONE*, *6*(10), e23223. https://doi.org/10.1371/journal.pone.0023223

Schug, J., Takagishi, H., Benech, C., & Okada, H. (2016). The development of theory of mind and positive and negative reciprocity in preschool children. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00888

Schurz, M., Aichhorn, M., Martin, A., & Perner, J. (2013). Common brain areas engaged in false belief reasoning and visual perspective taking: a meta-analysis of functional brain imaging studies. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00712

Scott, R. M., & Baillargeon, R. (2017). Early False-Belief Understanding. *Trends in Cognitive Sciences*, *21*(4), 237–249. https://doi.org/10.1016/j.tics.2017.01.012

Shaw, A., Montinari, N., Piovesan, M., Olson, K. R., Gino, F., & Norton, M. I. (2014). Children develop a veil of fairness. *Journal of Experimental Psychology: General*, *143*(1), 363–375. https://doi.org/10.1037/a0031247

Sheskin, M., Bloom, P., & Wynn, K. (2014). Anti-equality: Social comparison in young children. *Cognition*, *130*(2), 152–156. https://doi.org/10.1016/j.cognition.2013.10.008

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience & Biobehavioral Reviews*, *30*(6), 855–863. https://doi.org/10.1016/j.neubiorev.2006.06.011

Singer, T. (2009). Understanding others: Brain mechanisms of theory of mind and empathy. *Neuroeconomics: Decision Making and the Brain*, 251–268.

Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness? *Psychological Science*, *23*(2), 196–204. https://doi.org/10.1177/0956797611422072

Smith, C. E., Blake, P. R., & Harris, P. L. (2013). I should but I won't: Why young children endorse norms of fair sharing but do not follow them. *PloS One*, *8*(3), e59510. https://doi.org/10.1371/journal.pone.0059510

Struch, N., & Schwartz, S. H. (1989). Intergroup aggression: Its predictors and distinctness from in-group bias. *Journal of Personality and Social Psychology*, *56*(3), 364–373.

Takagishi, H., Koizumi, M., Fujii, T., Schug, J., Kameshima, S., & Yamagishi, T. (2014). The role of cognitive and emotional perspective taking in economic decision making in the Ultimatum Game. *PLoS ONE*, *9*(9), e108462. https://doi.org/10.1371/journal.pone.0108462

Thornton, A., & McAuliffe, K. (2015). Cognitive consequences of cooperative breeding? A critical appraisal: Cognitive consequences of cooperative breeding. *Journal of Zoology*, *295*(1), 12–22. https://doi.org/10.1111/jzo.12198

Todd, A. R., & Galinsky, A. D. (2014). Perspective-Taking as a Strategy for Improving Intergroup Relations: Evidence, Mechanisms, and Qualifications: Perspective-Taking and Intergroup Relations. *Social and Personality Psychology Compass*, *8*(7), 374–387. https://doi.org/10.1111/spc3.12116

Tsoi, L., Dungan, J. A., Chakroff, A., & Young, L. L. (2018). Neural substrates for moral judgments of psychological versus physical harm. *Social Cognitive and Affective Neuroscience*. https://doi.org/10.1093/scan/nsy029

Tsoi, L., Dungan, J., Waytz, A., & Young, L. (2016). Distinct neural patterns of social cognition for cooperation versus competition. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2016.04.069

Tsoi, L., Hamlin, J. K., Baron, A. S., Waytz, A., & Young, L. (under review). Developmental evidence for enhanced theory of mind for competitive versus cooperative contexts.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*(3), 383–403. https://doi.org/10.1037/0033-2909.134.3.383

Valdesolo, P., & DeSteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion*, *11*(2), 262–266. https://doi.org/10.1037/a0021302

Wagner, D. D., Kelley, W. M., & Heatherton, T. F. (2011). Individual differences in the spontaneous recruitment of brain regions supporting mental state understanding when viewing natural social scenes. *Cerebral Cortex*, *21*(12), 2788–2796. https://doi.org/10.1093/cercor/bhr074

Waytz, A., Epley, N., & Cacioppo, J. T. (2010). Social Cognition Unbound: Insights Into Anthropomorphism and Dehumanization. *Current Directions in Psychological Science*, *19*(1), 58–62. https://doi.org/10.1177/0963721409359302

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, *14*(8), 383–388. https://doi.org/10.1016/j.tics.2010.05.006

Waytz, A., Morewedge, C. K., Epley, N., Monteleone, G., Gao, J.-H., & Cacioppo, J. T. (2010a). Making sense by making sentient: Effectance motivation increases anthropomorphism. *Journal of Personality and Social Psychology*, *99*(3), 410–435. https://doi.org/10.1037/a0020240

Waytz, A., Morewedge, C. K., Epley, N., Monteleone, G., Gao, J.-H., & Cacioppo, J. T. (2010b). Making sense by making sentient: Effectance motivation increases anthropomorphism. *Journal of Personality and Social Psychology*, *99*(3), 410–435. https://doi.org/10.1037/a0020240

Waytz, A., & Young, L. (2014). Two motivations for two dimensions of mind. *Journal of Experimental Social Psychology*, *55*, 278–283. https://doi.org/10.1016/j.jesp.2014.08.001

Wellman, H. M., & Banerjee, M. (1991). Mind and emotion: Children's understanding of the emotional consequences of beliefs and desires. *British Journal of Developmental Psychology*, *9*(2), 191–214. https://doi.org/10.1111/j.2044-835X.1991.tb00871.x

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655–684.

Wheatley, T., Milleville, S. C., & Martin, A. (2007). Understanding animate agents: Distinct roles for the social network and mirror system. *Psychological Science*, *18*(6), 469–474. https://doi.org/10.1111/j.1467-9280.2007.01923.x

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*(5), 297–333. https://doi.org/10.1037/h0040934

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and Cooperation. *Psychological Science*, *20*(1), 1–5. https://doi.org/10.1111/j.1467-9280.2008.02253.x

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128.

Woodruff, G., & Premack, D. (1979). Intentional communication in the chimpanzee: The development of deception. *Cognition*, *7*(4), 333–362. https://doi.org/10.1016/0010-0277(79)90021-0

Yirmiya, N., Erel, O., Shaked, M., & Solomonica-Levi, D. (1998). Meta-analyses comparing theory of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. *Psychological Bulletin*, *124*(3), 283–307.

Yoder, K. J., & Decety, J. (2014). The Good, the Bad, and the Just: Justice sensitivity predicts neural response during moral evaluation of actions performed by others. *Journal of Neuroscience*, *34*(12), 4161–4166. https://doi.org/10.1523/JNEUROSCI.4648-13.2014

Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the*

National Academy of Sciences, *104*(20), 8235–8240. https://doi.org/10.1073/pnas.0701408104

Young, Liane, & Saxe, R. (2009a). An fMRI investigation of spontaneous mental state inference for moral judgment. *Journal of Cognitive Neuroscience*, *21*(7), 1396–1405. https://doi.org/10.1162/jocn.2009.21137

Young, Liane, & Saxe, R. (2009b). Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia*, *47*(10), 2065–2072. https://doi.org/10.1016/j.neuropsychologia.2009.03.020

Zaki, J., & Ochsner, K. (2016). Empathy. In *Handbook of Emotion* (4th ed., pp. 871–884). New York: Guilford.

3.6 APPENDIX

Format:

Background Information
Condition - Cooperation
Condition - Competition
Condition - Neutral
True/False Statement
Answer to True/False

Sally put her ball in the box and left the room. Anne came in to the room, moved the ball from the box to the cabinet, and left the room.
Anne wanted to help Sally by putting the ball in the right place.
Anne wanted to play a trick on Sally by moving the ball.
Anne wanted to leave the room to go play in the park.
When Sally comes back, she will think that her ball is in the box.
TRUE

Lucy put two piles of clothes on her bed: one to keep and one for donation. While Lucy was in the attic to look for more clothes, Christina took the pile Lucy wanted to keep and brought it downstairs.
Christina thought the pile of clothes she took was meant for donation.
Christina thought the pile had some really nice clothes she would want for herself.
Christina thought the pile contained about twenty shirts.
When Lucy returns, she will think that her keep pile is still on her bed.
TRUE

Roxanne went to bed, leaving an unfinished puzzle in the living room so that she could solve it the next day. Max saw the puzzle and solved it that night.
Max wanted to help Roxanne solve the puzzle.
Max wanted to show Roxanne that he beat her to solving the puzzle.
Max wanted to finish the puzzle even if it took the whole night.
When Roxanne wakes up, she will think the puzzle is completed.
FALSE

Brad spent over 12 hours in the communal living area watching DVDs he borrowed from the library, and he planned on rewatching some of them the next day. While he was sleeping, his roommate Tracy returned all the DVDs to the library.
Tracy thought Brad was done with the DVDs and did it as a friendly gesture.
Tracy thought that she could finally have the living room to herself now that the DVDs were gone.
Tracy thought it was interesting that all the movies were filmed in black and white.
When Brad wakes up, he will think that the DVDs are still at the house.
TRUE

Davon cleared his desk and put all his items on the floor because he needed more desk space to do his work. When he left the house to run some errands, Nikki put all the items back on the table.
Nikki wanted to keep the floor clean so no one would trip on anything.
Nikki wanted to disrupt Davon's work so that he would be further behind.
Nikki wanted to borrow some of the items from Davon later.
When Davon returns, he will think that his desk is cleared of any items.
TRUE

Trevor went to the pantry to get spices to make his soup less bland. While Trevor was away, Lauren, the instructor of the culinary class, added unlabeled spice to Trevor's soup.
Lauren wanted to improve the flavors of the bland soup.
Lauren wanted to sabotage Trevor by adding unknown spice.
Lauren wanted to look at every student's soup.
When Trevor gets back to his workstation, he will think that his soup needs more spice.
TRUE

Paula forgot the combination to her lock, so she left her locker unlocked and went in to class. Anthony saw the unlocked locker and locked it.
Anthony wanted to prevent people from stealing Paula's things.
Anthony wanted to prevent Paula from getting to her belongings.
Anthony wanted to get to his next class on time so he walked briskly.
When Paula gets back, she will think that the locker is locked.
FALSE

Stan needed to get to his fencing match soon, so he asked Duncan to request an Uber while he was getting ready to leave. Instead of requesting an Uber, Duncan requested a Lyft.
Duncan wanted Stan to get to his match soon and Lyft had an earlier arrival time than Uber.
Duncan wanted Stan to get to his match late and Lyft had a later arrival time.
Duncan wanted to use a promo code for a free ride.

When Stan leaves the house, he will think an Uber will be there.
TRUE

Amy was told that the results of her medical tests would come in the mail that week. The mail arrived in a timely manner, but her daughter decided to hide the results from Amy for another two weeks.
Amy's daughter thought Amy had enough stress to deal with.
Amy's daughter thought that holding on to the results would prevent Amy from receiving treatment.
Amy's daughter thought the results were difficult to interpret.
When the week passes by, Amy will think that she has already received the mail.
FALSE

Parker moved the ladder close to the air vent and crawled into the vent. Asher moved the ladder from under the vent back to where it typically is.
Asher thought the ladder was there because someone forgot to return it back.
Asher thought playing a trick on Parker would be funny.
Asher thought it was hard not to scratch the floor when dragging the ladder.
When Parker is ready to come back down, he will think the ladder is right under the vent.
TRUE

Stacy went to get some water after her workout at the treadmill, leaving behind her towel. Elena, who wanted to use the treadmill, placed the towel on a nearby elliptical.
Elena thought that placing the towel there would make it easy for Stacy to find her towel.
Elena thought moving Stacy's towel would annoy Stacy.
Elena thought that she stayed at the gym long enough so she left soon after.
When Stacy gets back, she will think that her towel is on the elliptical.
FALSE

Ashley chopped some vegetables to use for cooking the next day. Her roommate Maggie saw the vegetables in the fridge and decided to cook them while Ashley was at work.
Maggie wanted to have Ashley's dinner ready for her when she returned.
Maggie wanted to shorten her cooking time by using Ashley's chopped vegetables.
Maggie wanted her meal to contain vegetables and meat.
When Ashley gets back home, she will think her vegetables are in the fridge.
TRUE

Chris was moving out of his furnished apartment and was packing up items in the kitchen. Without any input from Chris, Andrew loaded the desk that came with the apartment into the truck.
Andrew wanted to help move bulkier items.

Andrew wanted to get Chris in trouble with the building superintendent.
Andrew wanted a Gatorade after moving the desk.
When Chris gets to the truck, he will think the desk is in the truck.
FALSE

Brittany decided to take golf lessons and planned on signing up for them later that day.
Sam signed up for the last spot not knowing that Brittany didn't sign up yet.
Sam thought that it would be great if he and Brittany took lessons together.
Sam thought that formal lessons would help him be better at golf than Brittany.
Sam thought the lessons were really cheap given the frequency of the lessons.
When Brittany checks the sign-up sheet, she will think that there are no spots left.
FALSE

Joe, who was running late that morning, dropped his bag by the entryway and ran to the kitchen to pick up something he forgot. While Joe was in the kitchen, Steve put the bag away.
Steve thought Joe had left and didn't want the bag to clutter up the entryway.
Steve thought Joe would get frustrated about losing his bag.
Steve thought Joe's bag was really heavy given its size.
When Joe returns from the kitchen, he will think his bag is in the entryway.
TRUE

Janice placed the heavy vase she had just purchased at the yard sale on the floor outside the public restroom. While Janice was in the restroom, Heather moved the vase to the vase section of the yard sale.
Heather thought the vase had been misplaced so she moved it there.
Heather thought doing this would really upset Janice.
Heather thought the vase section of the yard sale contained a nice variety of different vases.
When Janice is done with the restroom, she will think her vase is at the vase section of the yard sale.
FALSE

Anna wanted to go mountain climbing so she bought some gear. Her husband, David, saw the gear on the table and brought it back to the store while Anna was at work.
David thought that Anna needed nicer gear for several safety reasons.
David thought that Anna doing mountain climbing would threaten his masculinity.
David thought the store was so disorganized that it was hard to find specific items.
When Anna comes home, she will think that her gear is on the table.
TRUE

Jeff wanted a nice photo for his dating profile, so he asked Allie to enhance a photo of him. Allie removed the background and added a plain white background.
Allie wanted to help create a decent photo for what she thought was for a passport.
Allie wanted to make Jeff's photo look boring and unappealing.
Allie wanted to make other alterations but ran short on time.
When Jeff gets the photo, he will think the photo will have a plain white background.
FALSE

Erica removed the tire on her bike and went to the bike shop to get a new replacement tire. While she was away, Brian saw the tire and put the tire back on the bike.
Brian wanted to help Erica out with something that he's good at doing.
Brian wanted Erica to struggle with getting her bike in working order.
Brian wanted to get a new bike like Erica's, but with bigger wheels and a lighter frame.
When Erica gets back, she will think the tire is removed from the bike.
TRUE

Brianna dislikes open umbrellas in indoor spaces, so she closed her wet umbrella and left it by the front door before heading to the living room. Amanda, who came in after Brianna, opened Brianna's umbrella and left it open by the front door.
Amanda thought opening it would get the umbrella to dry quicker.
Amanda thought opening the umbrella would annoy Brianna.
Amanda thought the blue and orange swirls on the umbrella nicely complemented each other.
When Brianna comes out to get the umbrella, she will think the umbrella is closed.
TRUE

Amanda found Mike's paint can in the closet and left it out to be thrown away since the color was ugly. While she was away, Mike saw the paint can and painted their bedroom with it.
Mike thought it was the can Amanda wanted him to use to paint their room.
Mike thought the color that he picked was more welcoming than the one Amanda had wanted.
Mike thought it would take 30 minutes to paint the room, but it actually took over two hours.
When Amanda gets to the bedroom, she will think the bedroom will be newly painted.
FALSE

Maria left all of her clothes out as she was packing for vacation and left to get more new clothes. Dani saw that the clothes were out and put them away.
Dani wanted to help Maria keep her room clean.
Dani wanted to make packing more difficult for Maria.
Dani wanted her sense of style to be as good as Maria's.

When Maria comes back with new clothes, she will think her clothes are packed away.
FALSE

Maya was working on her portion of a complex multi-person project and took a break by taking a short nap. While Maya was napping, Brianna finished the project.
Brianna wanted to help Maya with the project.
Brianna wanted more credit for the work.
Brianna wanted to take a nap, too.
When Maya wakes up, she will think the project is finished.
FALSE

Juan checked his calendar and saw that his schedule was free at noon. While he was making plans to eat out for lunch, his secretary scheduled a meeting for Juan at noon.
His secretary thought this was a good business opportunity for Juan.
His secretary thought Juan would get mad for having to skip lunch yet again.
His secretary thought the meeting should take place in Meeting Room B.
When Juan checks his calendar again, he will think that there is a meeting at noon.
FALSE

Kate, prior to chorus practice, came up with a song and showed it to Grace. While Kate was away, Grace memorized it and showed it to their teacher.
Grace wanted to get advice from the teacher on how to make Kate's song even better.
Grace wanted to take the credit for coming up with the song.
Grace wanted to hit all the high notes without straining her voice.
When Kate gets back, she will think the teacher has already heard her song.
FALSE

Carla was baking a cake and asked Jane for vegetable oil. While Carla was looking for other ingredients in the pantry, Jane added sesame oil to the cake mix.
Jane thought there was no difference between sesame oil and vegetable oil.
Jane thought adding sesame oil to the mix would make Carla's cake taste weird.
Jane thought the cake should be done in 40 minutes.
When Carla gets back from the pantry, she will think the cake batter contains sesame oil.
FALSE

Caroline works as a fry cook flipping hamburgers. Her co-worker, Ali, saw that Caroline walked away from the grill and took the burger off the grill.
Ali wanted to prevent the burger from burning.
Ali wanted Caroline to get in trouble for walking away from the grill.
Ali wanted to make sure she didn't burn her fingers while moving the burger.
When Caroline returns, she will think the burger is on the grill.

TRUE

Jackson left his keys at home, so he kept his office door unlocked and open. While he was at a meeting, the janitor closed and locked the office door.
The janitor thought it was important to keep everything in the office safe.
The janitor thought locking the door would make it harder for Jackson to get into his office.
The janitor thought the office door plaque needed to be updated.
When Jackson returns to his office, he will think his door is unlocked.
TRUE

Jenny found a nice spot at the party to put down her beer. While she was getting food, Evan moved her beer from the windowsill to the floor.
Evan thought the wind might knock down the beer from the windowsill.
Evan thought his beer should be the one by the windowsill instead.
Evan thought there was no way the house would comfortably fit everyone.
When Jenny returns with her food, she will think that her beer is on the windowsill.
TRUE

Nicole put down twenty seashells she found at the beach and planned to pick them up after a swim. Jason moved the seashells further away from the water.
Jason wanted to keep the seashells from getting washed away.
Jason wanted to make Nicole think the seashells got washed away.
Jason wanted to lie on the sand and remain far from the water.
When Nicole gets back from the swim, she will think the seashells are where she dropped them.
TRUE

3.7 SUPPLEMENTARY MATERIAL

3.7.1 Theory of mind localizer

Table S1. Mean peak coordinates (in MNI) for each ROI defined using the functional localizer

	x	y	z	# of voxels	# of people (out of 34)
RTPJ	51.3	-54.6	17.1	79.6	30
LTPJ	-50.1	-58.9	20.4	65.3	29
PC	1.2	-59.5	27.5	74.6	28
DMPFC	1	52.4	26.8	51.2	21

4.0 Paper 3: Individual differences in theory of mind predict inequity aversion in children

Early in human development, children react negatively to receiving less than others, and only later do they show a similar aversion to receiving more. We investigated whether theory of mind can account for this developmental shift we see in middle childhood. To investigate this question, we conducted a face-to-face fairness task that involved a theory of mind manipulation, measured individual differences in theory of mind, and collected parent-ratings of their children's empathy, a construct related to theory of mind. We find that greater theory of mind capacities lead to more rejections of unequal offers, regardless of the direction of inequality, demonstrating that children with greater theory of mind are more likely to engage in costly compliance with fairness norms. These findings contribute to our growing understanding that people's concerns for fairness rely not just on their own thoughts and beliefs but on the thoughts, beliefs, and expectations of others.

This paper is co-authored with Katherine McAuliffe.

4.1 INTRODUCTION

Humans care deeply about fairness. Concerns about fairness are so strong that people are willing to sacrifice personal gain to avoid unfair resource distributions (Dawes, Fowler, Johnson, McElreath, & Smirnov, 2007; Fehr & Schmidt, 1999; Loewenstein, Thompson, & Bazerman, 1989). Indeed, people are not only willing to pay personal costs to avoid receiving less than others (*disadvantageous inequity aversion*), they also pay to avoid receiving more than others (*advantageous inequity aversion*). A strong aversion to inequality likely drives behavior across a range of contexts, such as punishment (Raihani & McAuliffe, 2012) and different forms of cooperation (Brosnan & Bshary, 2016), yet little is known about the cognitive mechanisms that give rise to it.

A developmental approach can offer key insights into the cognitive mechanisms supporting inequity aversion in humans since we can capitalize on the fact that there are distinct developmental trajectories for aversion to the two forms of inequity. A typical method for studying this developmental question is by using economic games, which provide researchers the ability to observe people's actual and often costly behaviors; this is an important consideration given known discrepancies between real and hypothesized behaviors (FeldmanHall, Dalgleish, et al., 2012; FeldmanHall, Mobbs, et al., 2012). Moreover, these games allow researchers to examine responses to inequity when all else is equal, eliminating concerns about other factors known to affect people's fairness behaviors (e.g., need or merit; Baumard, Mascaro, & Chevallier, 2012; Hamann, Bender, & Tomasello, 2014; Kanngiesser & Warneken, 2012; Rizzo & Killen, 2016). Results using these games show that four- to seven-year-olds reject allocations that put them at a disadvantage relative to their peer but accept allocations that will put them at an advantage, revealing a preference for relative advantage among this age group (Blake, McAuliffe, & Warneken, 2014; Sheskin, Bloom, & Wynn, 2014). Indeed, rates of rejection of advantageous inequity among this age group are so low, they do not significantly differ from rates of rejection of equal allocations. It isn't until children reach the age of 8 (and potentially later in non-Western cultures) that they start becoming averse to advantageous inequity (Blake et al., 2015; Blake & McAuliffe, 2011; McAuliffe, Blake, Kim, Wrangham, & Warneken, 2013; McAuliffe, Blake, Steinbeis, & Warneken, 2017). Why do we see this developmental difference between advantageous and disadvantageous inequity aversion?

A candidate mechanism for explaining the developmental dissociation between aversion to disadvantageous and advantageous inequity is theory of mind. That is, the capacity to consider the minds of others (e.g., mental states such as desires, beliefs, knowledge) may encourage children to respond negatively when others receive less than their fair share of resources relative to themselves. While this hypothesis has not been tested directly, related work has provided evidence linking consideration of other people's minds and viewpoints with improvements in interpersonal relations among adults (Bruneau & Saxe, 2012; Galinsky & Moskowitz, 2000; Galinsky, Maddux, Gilin, & White, 2008; Todd & Galinsky, 2014) and with fairness behaviors in children (Schug, Takagishi, Benech, & Okada, 2016; Takagishi et al., 2014; but see Liu et al., 2016; Rochat et al., 2009). Notably, most of the developmental studies have targeted children at the preschool age (between 3- and 5-years-old) and have focused primarily on children's understanding of false beliefs, just one of many aspects of theory of mind (Bloom & German, 2000). Given that theory of mind continues to develop past the preschool age

(Flavell, 2004; Hughes & Leekam, 2004) and changes in fairness behaviors can be found in middle childhood, this line of work may not adequately capture the potential role theory of mind may play in promoting developmental changes in fairness behaviors. Moreover, to understand how theory of mind contributes to fairness specifically, we must steer away from a reliance on unilateral sharing contexts like the Dictator Game and complex two-player tasks like the Ultimatum Game (a commonly used game in fairness studies) since these tasks are unable to disentangle different motivations for children's behaviors. For instance, the Dictator Game assesses generosity, altruism, and fairness while the Ultimatum Game assesses strategic responses in the proposer and punishing behaviors (if faced with an unfair offer) in the recipient. Instead, we turn to the Inequity Game (Blake & McAuliffe, 2011), which specifically measures children's responses to distributional unfairness. Thus, using this game we can examine whether changes in theory of mind can account for the developmental shift we see in middle childhood between caring only about fairness for the self (disadvantageous inequity aversion) toward a more general concern for fairness not just for the self but for others as well (advantageous inequity aversion).

Here we test the idea that maturing theory of mind abilities can explain the developmental dissociation between aversion to disadvantageous inequity and advantageous inequity. In the main experimental task with the Inequity Game, the experimenter allocated resources (e.g., candy) equally or unequally to two children. One child made the decision on behalf of both children across all trials; he/she either accepted an allocation (in which case, each child received the amount that was allocated to him/her) or rejected an allocation (in which case, neither child received anything). The experimenter introduced the experimental task by focusing either on the knowledge states of players on both sides of the task apparatus (theory of mind condition) or on the physical attributes of the apparatus (control condition). We chose this manipulation instead of instructing children to think about the mental states of their partner so that we could examine the effect of changing children's spontaneous tendencies to mentalize on inequity aversion.

After the main experimental task (Figure 1a), the child making all the decisions also completed a theory of mind individual difference measure. Because we were interested in testing 6-9 year-olds, we were unable to use traditional theory of mind tasks, given that they typically target preschool children (e.g., Wellman, Cross, & Watson, 2001). Instead, we administered the child version of the Reading the Mind in the Eyes Test (Figure 1b; Baron-Cohen et al., 2001), a task developed by the same group who created the widely used and validated task for measuring individual differences in theory of mind in adults (Baron-Cohen et al., 2001). Additionally, we wanted to explore the relationship between inequity aversion and empathy, a construct related to but distinct from theory of mind (Singer, 2006, 2009). Both theory of mind and empathy support moral cognition (Bzdok et al., 2012) but each is subserved by different neural circuitry. Moreover, the constructs appear to be separate: while theory of mind is thought of as the ability to understand people's beliefs and thoughts, empathy is typically defined as the ability to understand people's feelings (but see Zaki & Ochsner, 2016). To assess children's empathy, we had parents complete a parent-rating of their child's empathy (Figure 1c; Dadds et al., 2008).

We made several predictions. First, we predicted that an aversion to disadvantageous inequity may not require a theory of mind, an idea supported by findings that rejections of this form of inequity take place even when children face such distributions in the absence of a social partner (McAuliffe et al., 2013). By contrast, aversion to advantageous inequity may result from children considering the mental states of their social partner; children may then weigh their partner's thoughts to receiving a relatively poor payoff with their own desire for the more attractive payoff. Following this reasoning, the relatively late emergence of aversion to advantageous inequity may thus be driven by developmental changes in theory of mind capacities in mid- to late-childhood. Because we were interested in examining how theory of mind contributes to advantageous and disadvantageous inequity aversion, the focus of our study was on two age groups: 6-7 year-olds, who show aversion to disadvantageous inequity but not to advantageous inequity, and 8-9 year-olds, who, in addition to being averse to disadvantageous inequity, start becoming averse to advantageous inequity. Our prediction was that children in the theory of mind condition (in which they considered another's mental state) would be more likely to reject advantageous inequity as compared to children in the control condition. Specifically, a theory of mind manipulation could potentially increase advantageous inequity aversion in one of two ways: (i) the manipulation could increase both younger and older children's propensity to consider what their peer might think and know, which in turn leads them to reject advantageous inequity, or (ii) the manipulation could increase younger but not older children's propensity to consider what their peer might think and know (because older children might already have that propensity, rendering the manipulation ineffective), leading us to see an effect of the manipulation for younger children but not for older children. With regard to disadvantageous inequity aversion, there are two possible outcomes: (i) a theory of mind manipulation could lead children to consider the mental states of their peer, making them less likely to reject inequity benefiting the peer, or (ii) a theory of mind manipulation could have no effect on rejections of disadvantageous inequity, especially if rejections of disadvantageous inequity are not driven by considerations of others' minds. Given the same line of reasoning as presented above for the main experimental task, we predicted that theory of mind, as captured by the individual difference measure, would positively correlate with rejections of advantageous inequity, and by contrast, negatively correlate or not correlate at all with rejections of disadvantageous inequity.

4.2 METHODS

4.2.1 Participants

Our final sample consisted of 123 pairs of children recruited from public parks in the [removed for blind review] area. Each pair was comprised of one decider (the person who made all the decisions) and one passive recipient. Twenty-one additional pairs participated in at least part of the study but were excluded due to invalid consent forms (3), experimenter error (6), discomfort (1), parental interference (1), insufficient understanding of English (2), because they explicitly told us that they had recently participated in a similar study (3), dislike of the resource used (Skittles) (4), and because the recipient repeatedly signaled to the decider using nonverbal cues (pointing) in an apparent attempt to influence their decisions (1). Additionally, a large number of children were excluded from our main analyses because their parents refused to fill out or did not completely fill out our empathy questionnaire (36) or because some children did not complete our theory of mind measure (3); however, the inclusion of these participants did not affect the main result of interest (see Supplementary Material).

Deciders were 6- to 9-years old, divided between two age groups: 6- and 7-year-olds (N = 62, 31 females, M = 83.4 months, range = 72-95 months) and 8- and 9-year-olds (N = 61, 30 females, M = 107.7 months, range = 96-119 months). Recipients paired with these deciders were typically within ± 2 years of the decider's age. Assignment to the decider role was first-come, first-serve with some exceptions: in our study, a small number of children were assigned to the recipient role because they were either outside the 6-9 years range from which we were recruiting deciders (5), had intently watched others play the game (1), could not consume the resource used in the study (e.g., the resource (Skittles) was not kosher; 1), or because a child of a different gender was needed at the end of data collection to keep the number of male and female deciders relatively even (1).

Our goal was to recruit approximately 20 deciders per cell (2 [Condition: theory of mind vs. control] x 2 [Inequity Type: advantageous vs. disadvantageous] x 2 [Decider Age Group: 6&7 vs. 8&9]). Our target sample size was based on previous publications with similar study designs [removed for blind review]. Note that our number of participants per cell is inexact (Table S1) because we recruited children opportunistically in public parks and had several exclusion criteria. Before testing, parents and children 8-years-old and older provided informed written consent and all children provided verbal assent. This study was approved by the [removed for blind review] Institutional Review Board.

4.2.2 Design

We employed a 2 (Inequity Type: advantageous or disadvantageous) x 2 (Condition: theory of mind or control) x 2 (Decider Age Group: 6&7 or 8&9) x 2 (Distribution: equal or unequal) design with Inequity Type, Condition, and Decider Age Group as between-participant factors and Distribution as a within-participant factor. Each pair of children received 12 trials: six trials with equal distributions (one candy for decider, one for recipient; 1-1) and six with unequal distributions (either advantageous, 4-1, or disadvantageous, 1-4; see Table S1 for sample breakdown). The order of equal and unequal trials for each participant was randomized.

4.2.3 Procedure

After obtaining consent, children were introduced to the Inequity Game (Blake & McAuliffe, 2011) and its associated apparatus. One child, the decider, had control of two handles (Figure 1a) and made all decisions in the game while the recipient was passive. Pairs of children were unfamiliar with one another.

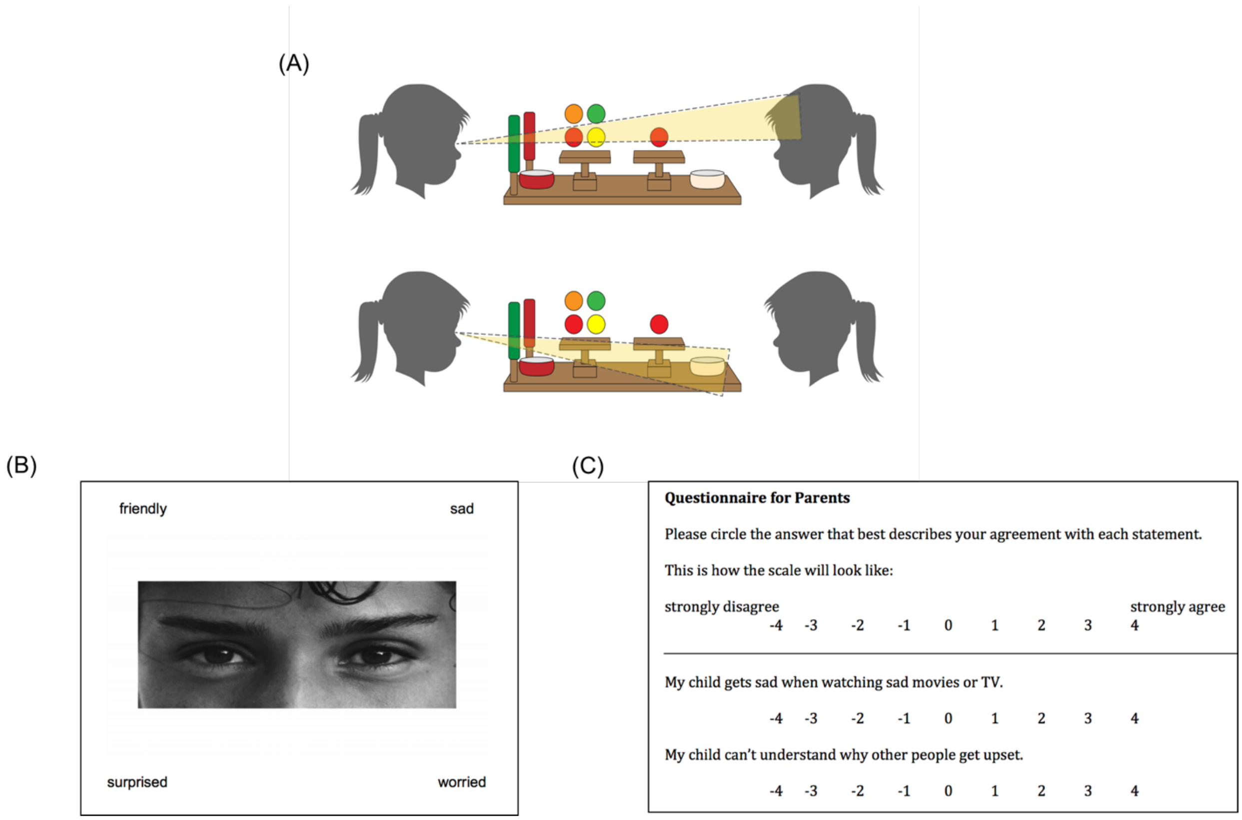


Figure 1. Experimental procedure. (A) Inequity Game (apparatus and participants are depicted). Top: *Theory of Mind* condition - participants focused on each other's mental states (i.e., knowledge states); Bottom: *Control* condition - participants focused on the physical attributes of the apparatus. (B) Child version of the Reading the Mind in the Eyes Test (RMET) – an individual difference measure of theory of mind. (C) Griffith Empathy Measure – a parent-rating measure of a child's empathy

To examine whether focusing on the knowledge states of both participants on both ends of the apparatus would influence children's responses to unfair distributions, children were assigned either to the *Theory of Mind* condition or the *Control* condition (Figure 1a). In the *Theory of Mind* condition, the experimenter introduced the setup by having children focus on the knowledge states of players on both sides of the apparatus. Children stood on the same side of the apparatus, and the experimenter said: "If you're on this side [pointing to the tray for the side closest to the children], you know what's on this side AND you know what's on that side [pointing to the other side's tray]". Children then switched to the other side of the apparatus and the same phrase was repeated by the experimenter. In the *Control* condition, an experimenter introduced the apparatus by having children focus on the physical attributes of the apparatus. Children stood on the same side of the apparatus, and the experimenter said: "If you're on this side [pointing to the tray for the side closest to the children], you can see that on this side, the bowl is [color of the bowl]". Children then switched to the other side of the apparatus and the same phrase was repeated by the experimenter with the appropriate color of the bowl (the color of the two bowls always differed).

Following this manipulation, the children took their seats. The decider sat on the side of the apparatus with the two handles used to make decisions in the game, and the recipient sat across from the decider. Throughout the entire experiment, the experimenter would place the Skittles on the recipient's tray before putting them on the decider's tray to ensure that the decider would pay attention to the recipient's payoff before attending to his or her own payoff.

After the experimenter introduced the game setup to the players, the experimenter demonstrated the use of both handles: the green handle was used to accept the distribution, and the red handle was used to reject the distribution. When the green handle was pulled, the trays tipped outward, causing the Skittles to fall into the two players' bowls. When the red handle was pulled, the trays tipped inward, causing the Skittles to fall into a middle bowl, in which case neither child got the Skittles.

After the demonstration, deciders were presented with three practice trials to ensure they understood the apparatus and the effects of pulling on both handles. The three practice trials contained an equal distribution (1 for the decider, 1 for the recipient; 1-1), an advantageous inequity distribution (1 for the decider, 0 for the recipient; 1-0), and a disadvantageous inequity distribution (0 for the decider, 1 for the recipient; 0-1). If the decider pulled the green handle for all three practice trials, the decider got an extra 1-1 trial and was asked to pull the handle that would make the candies fall into the middle bowl; if the decider pulled the red handle for all three practice trials, the decider got an extra 1-1 trial and was asked to pull the handle that would make the candies fall into the bowls on the side. The 1-1 practice trial was always administered first and then the order of the other two trials was counterbalanced between participants.

After the deciders completed the practice trials, the experimenter asked both players comprehension questions (questions were asked to the recipient first)[2]. In the *Theory of Mind* condition, the experimenter asked each child two questions: "You know how much candy is on your side; do you know how much candy is on [name of the other child]'s side? Does [name of the other child] know how much candy is on your side?". If a child could not correctly respond spontaneously to a question, the question was repeated with gestures to the two trays. In the *Control* condition, the experimenter asked each child: "You have a [color of the bowl] bowl; what color is [name of the other child]'s bowl?". If the child still could not answer the question after being asked 3 times, the experimenter would state the answer[3]. Children were then instructed to stay quiet during the trials.

---

[2] A subset of participants (N = 33) were asked comprehension questions before the practice trials. Doing this did not qualitatively affect deciders' decisions in the game.
[3] Most children correctly answered the questions in their first try. In the *Theory of Mind* condition, one participant correctly answered the first question on the 2nd try; two participants correctly answered the second question on the 2nd try. In the *Control* condition, one participant answered the question correctly on the 3rd try.

While children were playing the game, the parent of the decider was asked to fill out the Griffith Empathy Measure (Figure 1c; Dadds et al., 2008). The Griffith Empathy Measure is a 23-item parent-report measure of a child's empathy. Statements, which raters had to assess their agreement with a 9-point Likert scale, ranged from "Your child becomes sad when other children are sad" to "Your child rarely understands why other people cry". After the game, deciders met a second experimenter and completed the Reading the Mind in the Eyes test (RMET) (Figure 1b, Baron-Cohen et al., 2001), an individual difference measure of theory of mind. This test has 28 items, with each item consisting of an image of someone's eyes. For each item, children had to respond with which of four words they thought would best match what the person is thinking or feeling.

4.2.4 Coding and Analyses

Data were coded live by the main experimenter. Additionally, we had consent to video record the majority of sessions (78.9%), which were coded by an independent research assistant. Reliability between live and video coding was very good (Cohen's Kappa = 0.947).

All statistical analyses were conducted in R (version 3.3.2; R Core Team, 2015). Decisions were analyzed using Generalized Linear Mixed Models (GLMMs) with a binary response term (accept or reject; reject = 1). Mixed models were run using the package 'lme4' (Bates, Mächler, Bolker, & Walker, 2015b).

We were primarily interested in understanding whether children's responses to disadvantageous and advantageous distributions depended on our theory of mind manipulation, the child's RMET score, and their parent-reported empathy score. In addition, we were interested in understanding whether our theory of mind manipulation differentially influenced children across our two age groups. To these ends, our full model included the following predictor variables: Distribution (equal vs. unequal), Inequity Type (advantageous vs. disadvantageous), Condition (theory of mind vs. control), Decider Age Group[4] (6- to 7-year-olds vs. 8- to 9-year-olds), Decider Gender (female vs. male), RMET Score (which was scaled), and Griffith Score (which was also scaled). We examined (1) the two-way interaction between Distribution and Condition, the two-way interaction with Inequity Type, and the four-way interaction with Decider Age Group, (2) the two- and three-way interactions between Distribution, Inequity Type, and RMET Score and (3) the two- and three-way interactions with Distribution, Inequity Type, and Griffith Score. We also included two-way interactions between Decider Age Group and each of the two individual difference measures to control for possible interaction effects related to age. Participant was included as a random effect, and we fit a random slope and intercept for each Participant because the full model with a random slope and intercept provided a significantly better fit to the data than the full model with just the random intercept (LRT, $\chi^2(2) = 81.52$, $p < 0.001$). By fitting a random intercept and a random slope for each participant, we allow the intercept and effect of distribution to vary across individuals. Because GLMMs do not allow for missing data in predictors, we limited our analyses to data from participants for whom we had decision data as well as RMET and Griffith scores (N = 123 out of a possible 159 pairs for whom we had decision data). However, the results regarding the RMET do not change when we include

[4] We also ran analyses using age as a continuous variable, and the same pattern of results held across all LRTs. Results are described in Supplementary Material.

participants for whom we only had decision data as well as RMET scores (N = 156 out of a possible 159 pairs for whom we had decision data; see Supplementary Material).

To assess the importance of our predictors of interest, we performed likelihood ratio tests (LRTs) to test whether the model including a given term would provide a better fit to the data than a model without that term. When plotting results, we used binomial confidence intervals, which were corrected according to the Agresti method (Agresti & Coull, 1998).

4.3 RESULTS

The output from a reduced model is included in the main text (Table 1). For the full model, please see Supplementary Material.

Table 1. Model output from a reduced model that includes all the two-way interactions of interest (Distribution x Condition, Distribution x RMET score, Distribution x Griffith score) (in bold). Estimates and standard errors (s.e.) of fixed effects in Generalized Linear Mixed Models predicting children's rejection behaviors. Baselines were set as follows: Distribution = equal, Inequity Type = AI, Condition = Control, Decider Age Group = 6&7, Decider Gender = female. Griffith score and RMET score were included as continuous predictors. Table also shows goodness-of-fit statistics.

	Model
Intercept	-2.66 (0.38)***
Distribution	1.39 (0.54)*
Inequity Type	-0.02 (0.31)
Condition	0.65 (0.32)*
Decider Age Group	0.01 (0.24)
RMET score	0.03 (0.20)
Griffith score	-0.16 (0.20)
Decider Gender	0.00 (0.23)
Distribution x Inequity Type	1.60 (0.57)**
Distribution x Condition	**-0.35 (0.56)**
Distribution x RMET score	**0.63 (0.28)***
Distribution x Griffith score	**0.04 (0.28)**
Decider Age Group x RMET score	0.04 (0.25)
Decider Age Group x Griffith score	0.35 (0.22)
AIC	1398.54
BIC	1488.59
Log Likelihood	-682.27
Num. obs. (trials)	1476
Num. groups: ID (Participants)	123
Variance: ID (Intercept)	1.06
Variance: ID (Distribution)	5.93
Covariance: ID (Intercept) (Distribution)	-1.81

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, †$p < 0.1$

4.3.1 Effects of the theory of mind manipulation

We first tested the effects of our theory of mind manipulation on children's decisions (Figure 2). Neither the four-way interaction between Distribution, Inequity Type, Condition, and Decider Age Group, the three-way interaction between Distribution, Inequity Type, and Condition, or the two-way interaction between Distribution and Condition were significant (four-way: LRT, $\chi^2(1) = 3.48$, $p = 0.06$; three-way: LRT, $\chi^2(1) = 2.87$, $p = 0.09$; two-way: LRT, $\chi^2(1) = 0.70$, $p = 0.40$). Entering age as a continuous variable did not change these patterns of results (see Supplementary Material). These results suggest the control and theory of mind conditions did not differentially affect the rate of rejection for equal and unequal trials.

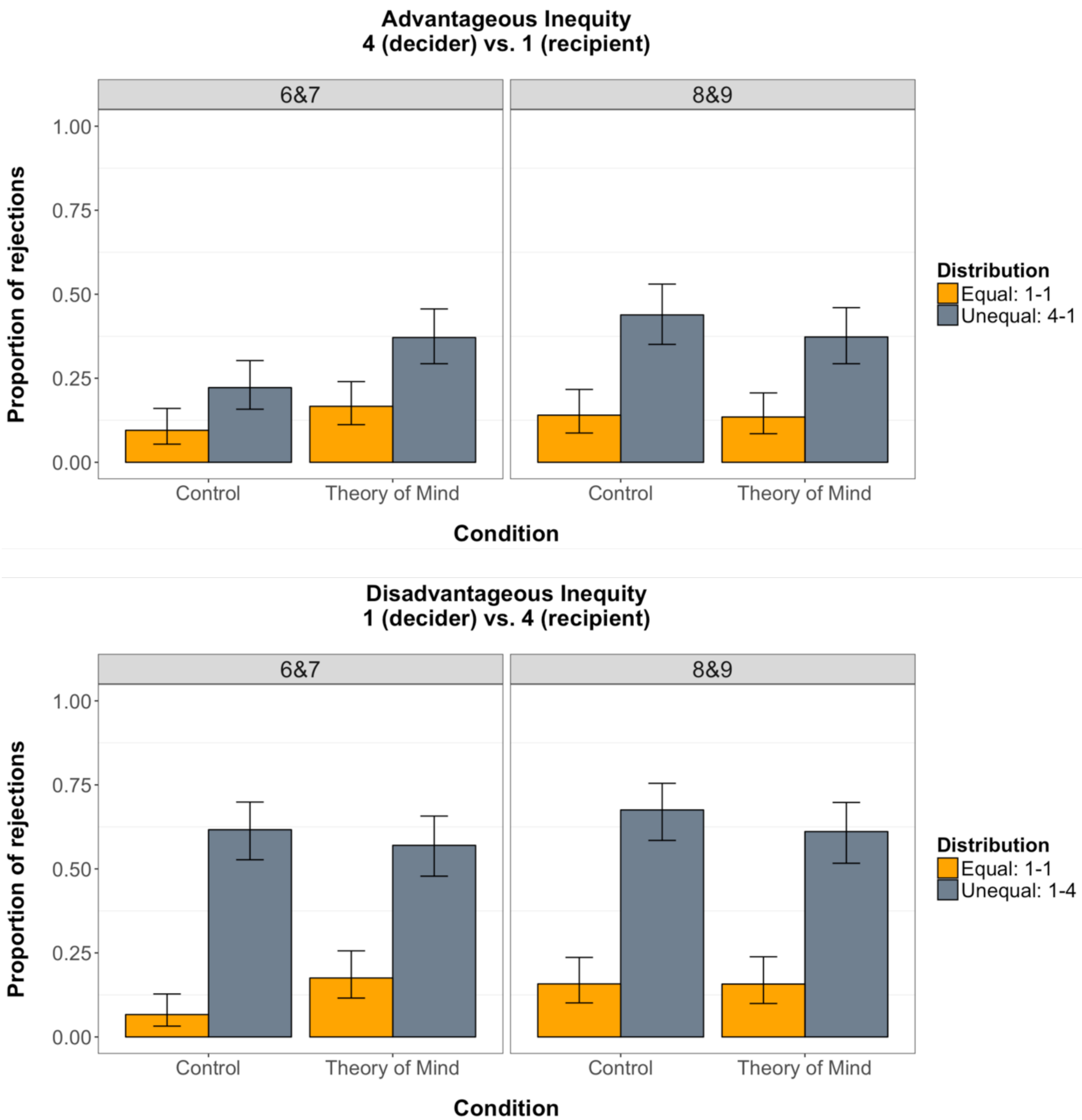


Figure 2. Rejections of equal and unequal distributions by Inequity Type (advantageous inequity [top] vs. disadvantageous inequity [bottom]), Condition (theory of mind vs. control), and Decider Age Group (6&7 vs. 8&9). Error bars denote 95% CI.

4.3.2 Relationship between inequity aversion and individual differences in theory of mind

We then examined whether theory of mind, as assessed by performance on the RMET, could predict children's decisions (Figure 3). While the three-way interaction between Distribution, Inequity Type, and RMET score was not significant (LRT, $\chi^2(1) = 0.30$, p = 0.58), the two-way interaction between Distribution and RMET score was significant (LRT, $\chi^2(1) = 4.45$, p = 0.035; model output can be found in Supplementary Material). This result suggests that while RMET score was not associated with rates of rejections of equal trials, RMET score was positively associated with rate of rejections of unequal trials, and this pattern was found for deciders faced with unequal trials that were advantageous as well as deciders faced with unequal trials that were disadvantageous[5]. The pattern is similar, albeit weaker, when we remove interactions with the Griffith score from the model so that the model is fit with data from a larger sample of participants (N = 156; for more details, see Supplementary Material).

---

[5] We were initially concerned that our theory of mind manipulation would affect performance on the RMET. To test this, we compared RMET scores for participants in the Theory of Mind vs. Control conditions. A Welch two-sample t-test confirmed that RMET scores did not differ between the two conditions ($M_{\text{Control}} = 16.73$; $M_{\text{Theory of Mind}} = 16.67$), $t(111.07) = 0.07$, $p = 0.94$.

4.3.3 Relationship between inequity aversion and individual differences in empathy

We also examined whether empathy, as assessed by the score on the Griffith Empathy Measure, could predict children's decisions (Figure 3). Neither the three-way interaction between distribution, inequity type, and Griffith score nor the two-way interaction between distribution and Griffith score were significant, although the three-way interaction was marginal (three-way: LRT, $\chi^2(1) = 2.66$, $p = 0.10$; two-way: LRT, $\chi^2(1) = 0.007$, $p = 0.93$). Entering age as a continuous variable in the model did not change the pattern of results (see Supplementary Material). These results provide some weak evidence that the interaction between inequity type and Griffith score differed for advantageous inequity and disadvantageous inequity.

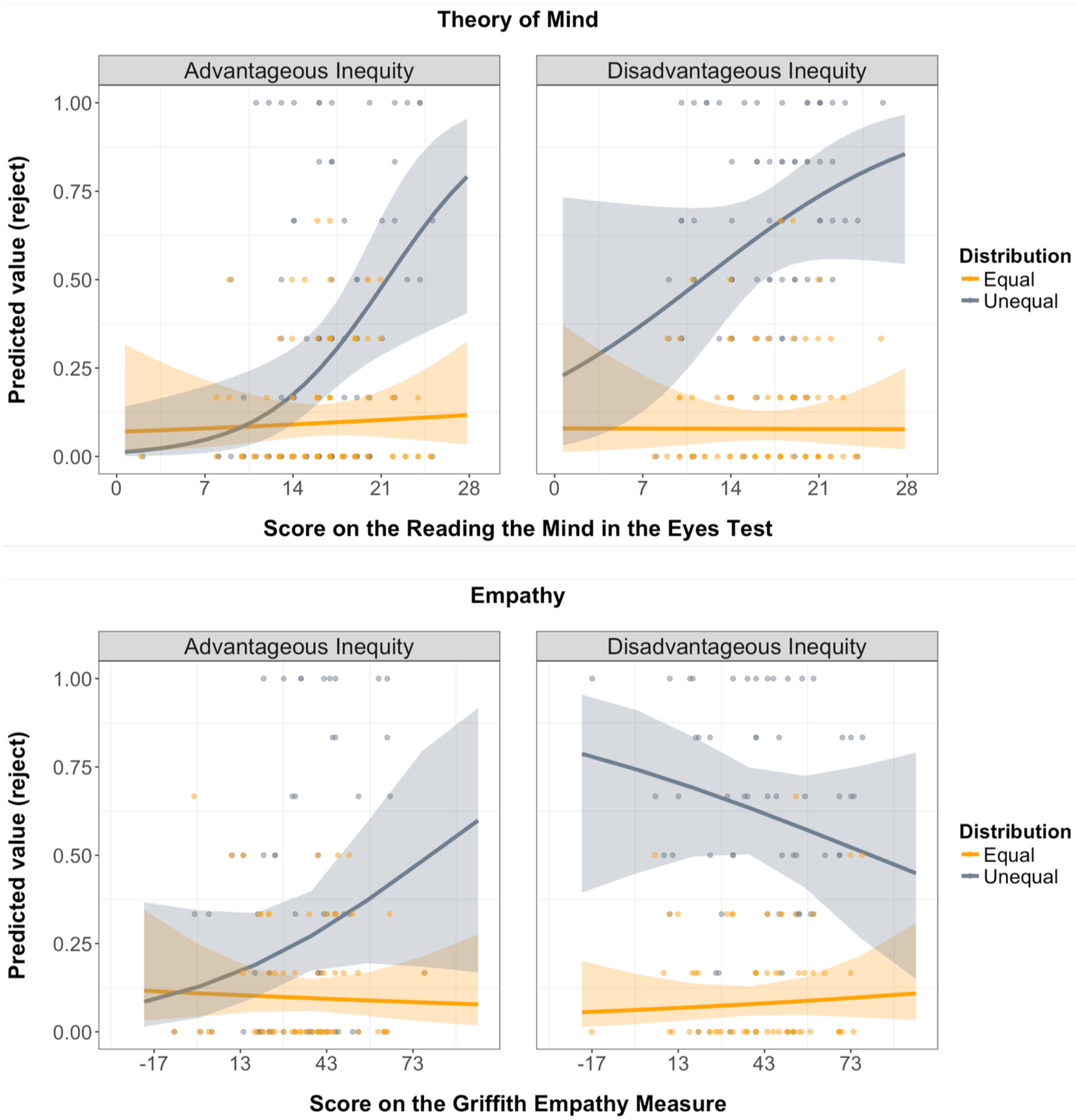


Figure 3. Rejections of equal and unequal distributions by each of the individual difference measures: Reading the Mind in the Eyes Test (top) and Griffith Empathy Measure (bottom). Lines and ribbons denote predicted values and 95% CIs extracted from the full model. Dots denote raw data.

4.4 DISCUSSION

Our aim was to understand whether theory of mind can account for children's emerging fairness preferences. Specifically, does an increased capacity to consider the minds of others explain why children shift from a self-focused response to unfairness (disadvantageous inequity aversion) to a more generalized concern for others with regard to fairness (advantageous inequity aversion)? We found that children with higher theory of mind scores (scores on the RMET) were more likely to reject unfair distributions than were those with lower scores. Interestingly, children with higher scores were more likely to reject inequity in both advantageous and disadvantageous directions. This finding suggests that a propensity to consider the mental states of others leads to costly compliance with fairness norms.

Our original hypothesis was that children with greater theory of mind capacities would be more likely to consider the mental states of their partner, leading them to (1) reject more unequal distributions that would benefit them more than their partner (advantageous inequity) and (2) potentially reject fewer unequal distributions that would benefit their partner more (disadvantageous inequity). Contrary to our hypotheses, higher theory of mind scores were associated with higher rates of rejections of unequal allocations regardless of the type of the inequity. That is, children with greater theory of mind capacities were more likely to enforce equality regardless of whether they or their partner would have benefited from the unequal resource division. One possible explanation for this counterintuitive result is that deciders with greater theory of mind capacities may be more adept at recognizing that the partner would expect them to uphold fairness norms and reject any form of unfairness. Violating fairness norms may cause the partner to think negatively of the decider, a possibility that may encourage deciders to reject both advantageous and disadvantageous forms of unfairness. By this account, theory of mind enables children to consider their partner's expectations, bringing their behavior in line with those expectations. This account could explain why, despite knowing that resources *should* be divided equally between people (with expectations of fair distributions developing as early as infancy; Geraci & Surian, 2011; Schmidt & Sommerville, 2011; Sloane, Baillargeon, & Premack, 2012), young children continue to allocate more resources to themselves than to others (Smith, Blake, & Harris, 2013). It is not until they reach middle childhood that children start to more consistently allocate resources fairly between themselves and others (Smith et al., 2013), a developmental change potentially facilitated by changes in their ability to consider the beliefs and expectations of others.

Work on how expectations affect behavior shows that children in our age range are able to adjust their behavior in line with others' expectations (Aloise-Young, 1993; Apfelbaum, Pauker, Ambady, Sommers, & Norton, 2008; Banerjee, 2002). Additionally, concerns for how they are perceived by others (i.e., reputational concerns) influence children's social behaviors (Engelmann & Rapp, 2018): for instance, school-age children recognize the need to behave differently to impress peers versus adults (Banerjee, 2002) and they are more generous when an audience is aware of their decisions (Buhrmester, Goldfarb, & Cantrell, 1992; Engelmann, Herrmann, & Tomasello, 2012; Leimgruber, Shaw, Santos, & Olson, 2012). In the context of fairness, 6- to 8-year-old children start to behave less fairly when they are able to do so and still maintain the appearance of being fair in front of an experimenter (Shaw et al., 2014). These reputational concerns are upheld in adulthood; indeed, adults are more fair and generous when others are aware of their behavior and are more reluctant to be fair when they can maintain the appearance of being fair without actually being fair (Andreoni & Bernheim, 2009; Batson, Thompson, & Chen, 2002; Dana, Weber, & Kuang, 2007; Reis & Gruzen, 1976). Overall, these results suggest that school-aged children's concerns about others' impressions and expectations of them may affect their fairness behaviors. Future work could explicitly test the idea that theory of mind is helping children adjust their behavior in line with others' expectations by introducing different expectations and examining how well theory of mind can predict behavioral adjustments to these different expectations. Another line of inquiry could involve delineating the role of theory of mind in understanding others' expectations of fairness versus understanding the consequences of violating those expectations.

In contrast to our finding that individual differences in theory of mind predicted increases in rejections to both forms of inequity, individual differences in empathy revealed a different pattern. While we cannot draw strong conclusions based on the marginal three-way interaction between Griffith score, Distribution, and Inequity Type, we want to make a few observations regarding this intriguing pattern of results. In the advantageous inequity condition, higher scores on the Griffith Empathy Measure led to higher rates of rejection for unequal trials but not equal trials, a pattern similar to that found for RMET scores. In the disadvantageous inequity condition, on the other hand, higher scores on Griffith Empathy Measure did not lead to higher rates of rejections for unequal trials—if anything, higher scores were linked to lower rates of rejections for unequal trials. These results suggest that children who are more empathic may be more likely to consider the emotional states of their partner, leading them to reject more unequal distributions that put their partner at a disadvantage and reject less unequal distributions that would put their partner at an advantage. An intriguing possibility is that inequity aversion is a specific social behavior that reveals a dissociation between theory of mind and empathy. The dissociation is evident when expectations of fairness and emotional reactions to unfairness are at odds with one another (e.g., partner expecting fairness but would be happy to receive more resources), as in the case of disadvantageous inequity. A partner's expectation of fairness and the partner's emotional reaction to unfairness may be weighted differently depending on the participant's capacity for theory of mind and empathy, respectively. This idea aligns well with prior work suggesting that theory of mind and empathy are related but distinct social constructs (Singer, 2009). Future studies may wish to examine the extent to which people base their decisions to accept/reject an unfair allocation on their partner's facial expressions as another method by which to disentangle those higher in empathy and those higher in theory of mind.

While one aim of our study was to examine individual differences in theory of mind and empathy and their relation to inequity aversion, another aim was to see whether children would display differential levels of aversion to inequity when put in a situation in which the experimenter discussed the mental states of people from both sides of the task apparatus versus the physical attributes of the apparatus. Specifically, we wanted to examine whether aversion to advantageous inequity, which emerges later in middle childhood, could be elicited in a younger age group with a theory of mind manipulation. The difference between rejections to equal and unequal trials in the case of advantageous inequity did not differ for the *Theory of Mind* and *Control* conditions in the younger age group, which suggests that our theory of mind manipulation did not succeed in eliciting aversion to advantageous inequity in younger children. However, another interpretation is plausible: younger children in our sample showed advantageous inequity aversion in both conditions; that is, we found higher rates of rejection for advantageously unequal versus equal trials in the younger age group for both the *Theory of Mind* and *Control* conditions. This finding contrasts with prior work revealing no difference in rates of rejection to advantageously unequal and equal trials in this younger age group (Blake & McAuliffe, 2011; McAuliffe et al., 2013). Seeing as we used the same apparatus and general procedure (aside from how we introduce the setup and the manipulation checks we ask) as in these prior studies, we do not believe the difference is due to anything other than the way we introduced the setup in both the *Control* and *Theory of Mind* conditions. Moreover, we ran an unrelated study with the Inequity Game in parallel to this one and found the typical baseline pattern [citation removed for blind review], which suggests that the difference we find in the current study is unlikely to be due to cohort effects. Instead, we speculate that our control condition was not an effective control condition: perhaps the simple act of instructing the decider and recipient to move together from one end of the experimental apparatus to the other in order to examine the physical attributes of the apparatus elicited more fair behavior (greater rejection of advantageous inequity) because it put children in a collaborative frame of mind. Indeed, prior work has shown that collaboration can increase sharing (Hamann, Warneken, Greenberg, & Tomasello, 2011) and aversion to advantageous inequity (Corbit, McAuliffe, Callaghan, Blake, & Warneken, 2017) in children, and even synchronous actions can increase cooperation among adults (Valdesolo & DeSteno, 2011; Wiltermuth & Heath, 2009). Given this possibility, we cannot yet rule out the notion that inequity aversion is impervious to theory of mind manipulations. Nevertheless, this speculation is tangential to the main message of the paper: that individual differences in theory of mind predict inequity aversion in children.

4.4.1 Limitations and Caveats

A strength of our study is that we are able to capture individual differences in theory of mind in school-aged children and examine their relationship to fairness behaviors; most prior work on children's theory of mind has largely focused on binary measures (e.g., passing/failing false belief tasks). Nevertheless, we recognize that we used just one individual difference measure of theory of mind. Understanding how the relationship between theory of mind and fairness can be measured with different scales is a fruitful area of future inquiry, particularly given recent efforts made to develop and validate measures of theory of mind appropriate for school-aged children (e.g., Devine & Hughes, 2016).

Moreover, there may be many possible paths to norm acquisition. While other work has examined the role of self-regulatory processes in fairness-related behaviors, such as following stated norms of giving (Blake et al., 2015, but see Smith 2013), we chose to focus on theory of mind and empathy. We do note, however, that the unique contributions of theory of mind are difficult to assess given the link between theory of mind and cognitive processes such as executive functioning. Indeed, the association between theory of mind and executive functioning appears to be robust even when controlling for factors such as IQ, at least in typically developing preschool children

(Carlson & Moses, 2001; Carlson et al., 2002; Carlson et al., 2004; Hughes, 1998; Perner & Lang, 1999; Sabbagh, Xu, Carlson, Moses, & Lee, 2006). More relevant is evidence showing that, among adults, performance on the Reading the Mind in the Eyes task is affected by interference with inhibitory processing (Bull, Phillips, & Conway, 2008). However, whether the relationship between theory of mind and executive functioning is just as robust in middle childhood is unclear, given mixed (and scant) evidence showing either no relationship between the two (Pellicano, 2007) or a small to moderate relationship (Austin, Groppe, & Elsner, 2014; Devine, White, Ensor, & Hughes, 2016) in typically developing school-aged children. More research on theory of mind and executive functioning in school-aged children may be crucial for understanding the large changes we examine in fairness behaviors.

4.5 CONCLUSION

Our study reveals that individual differences in theory of mind in school-aged children predict enforcement to a norm of fairness. Contrary to our expectations, greater theory of mind predicts rejections of both forms of inequality: children with higher theory of mind are more likely to reject inequity both when they could benefit from it as well as when others could benefit from it. Our findings are consistent with the idea that children with greater theory of mind capacities are more attuned to their partner's expectation that fairness norms ought to be upheld, thus leading them to reject both forms of unfairness. These findings contribute to our growing understanding that people's concerns for fairness rely not only on their own thoughts and beliefs but also on the thoughts, beliefs, and expectations of others.

4.6 REFERENCES

Agresti, A., & Coull, B. A. (1998). Approximate is better than “exact” for interval estimation of binomial proportions. *The American Statistician*, *52*(2), 119–126.

Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Do visual perspective tasks need theory of mind? *NeuroImage*, *30*(3), 1059–1068. https://doi.org/10.1016/j.neuroimage.2005.10.026

Aloise-Young, P. A. (1993). The development of self-presentation: Self-promotion in 6- to 10-year-old children. *Social Cognition*, *11*(2), 201–222. https://doi.org/10.1521/soco.1993.11.2.201

Andreoni, J., & Bernheim, B. D. (2009). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, *77*(5), 1607–1636. https://doi.org/10.3982/ECTA7384

Apfelbaum, E. P., Pauker, K., Ambady, N., Sommers, S. R., & Norton, M. I. (2008). Learning (not) to talk about race: When older children underperform in social categorization. *Developmental Psychology*, *44*(5), 1513–1518. https://doi.org/10.1037/a0012835

Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: A 1-year longitudinal perspective. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00655

Banerjee, R. (2002). Audience effects on self-presentation in childhood. *Social Development*, *11*(4), 487–507. https://doi.org/10.1111/1467-9507.00212

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*(2), 241–251.

Baron-Cohen, Simon. (2000). Theory of mind and autism: A fifteen year review. *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience (2nd Ed.)*, 3–20.

Baron-Cohen, Simon, Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, *21*(1), 37–46.

Baron-Cohen, Simon, Wheelwright, S., Spong, A., Scahill, V., & Lawson, J. (2001). Are intuitive physics and intuitive psychology independent? A test with children with Asperger Syndrome. *Journal of Developmental and Learning Disorders*, *5*(1), 47–78.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015a). Fitting linear mixed-effects models using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015b). Fitting linear mixed-effects models using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Batson, C. D., Thompson, E. R., & Chen, H. (2002). Moral hypocrisy: Addressing some alternatives. *Journal of Personality and Social Psychology*, *83*(2), 330–339. https://doi.org/10.1037/0022-3514.83.2.330

Baumard, N., Mascaro, O., & Chevallier, C. (2012). Preschoolers are able to take merit into account when distributing goods. *Developmental Psychology*, *48*(2), 492–498. https://doi.org/10.1037/a0026598

Begeer, S., Rieffe, C., Terwogt, M. M., & Stockmann, L. (2003). Theory of mind–based action in children from the autism spectrum. *Journal of Autism and Developmental Disorders*, *33*(5), 479–487.

Blake, P. R., McAuliffe, K., Corbit, J., Callaghan, T. C., Barry, O., Bowie, A., … Warneken, F. (2015). The ontogeny of fairness in seven societies. *Nature*, *528*(7581), 258–261. https://doi.org/10.1038/nature15703

Blake, Peter R., & McAuliffe, K. (2011). "I had so much it didn't seem fair": Eight-year-olds reject two forms of inequity. *Cognition*, *120*(2), 215–224. https://doi.org/10.1016/j.cognition.2011.04.006

Blake, Peter R., McAuliffe, K., & Warneken, F. (2014). The developmental origins of fairness: the knowledge–behavior gap. *Trends in Cognitive Sciences*, *18*(11), 559–561. https://doi.org/10.1016/j.tics.2014.08.003

Blake, Peter R., Piovesan, M., Montinari, N., Warneken, F., & Gino, F. (2015). Prosocial norms in the classroom: The role of self-regulation in following norms of giving. *Journal of Economic Behavior & Organization*, *115*, 18–29. https://doi.org/10.1016/j.jebo.2014.10.004

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, *77*(1), B25-31.

Brosnan, S. F., & Bshary, R. (2016). On potential links between inequity aversion and the structure of interactions for the evolution of cooperation. *Behaviour*, *153*(9–11), 1267–1292. https://doi.org/10.1163/1568539X-00003355

Bruneau, E. G., & Saxe, R. (2012). The power of being heard: The benefits of 'perspective-giving' in the context of intergroup conflict. *Journal of Experimental Social Psychology*, *48*(4), 855–866. https://doi.org/10.1016/j.jesp.2012.02.017

Buhrmester, D., Goldfarb, J., & Cantrell, D. (1992). Self-presentation when sharing with friends and nonfriends. *The Journal of Early Adolescence*, *12*(1), 61–79.

Bull, R., Phillips, L. H., & Conway, C. A. (2008). The role of control functions in mentalizing: Dual-task studies of Theory of Mind and executive function. *Cognition*, *107*(2), 663–672. https://doi.org/10.1016/j.cognition.2007.07.015

Burkart, J. M., Hrdy, S. B., & Van Schaik, C. P. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, *18*(5), 175–186. https://doi.org/10.1002/evan.20222

Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language*, *28*(5), 606–637. https://doi.org/10.1111/mila.12036

Byrne, R. W., & Corp, N. (2004). Neocortex size predicts deception rate in primates. *Proceedings of the Royal Society B: Biological Sciences*, *271*(1549), 1693–1699. https://doi.org/10.1098/rspb.2004.2780

Byrne, Richard W., & Whiten, A. (Eds.). (1988). *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford : New York: Clarendon Press ; Oxford University Press.

Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., & Eickhoff, S. B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure and Function*, *217*(4), 783–796. https://doi.org/10.1007/s00429-012-0380-y

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*(4), 1032–1053.

Carlson, Stephanie M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development*, *11*(2), 73–92. https://doi.org/10.1002/icd.298

Carlson, Stephanie M, Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, *87*(4), 299–319. https://doi.org/10.1016/j.jecp.2004.01.002

Carruthers, P. (2016). Two Systems for Mindreading? *Review of Philosophy and Psychology*, *7*(1), 141–162. https://doi.org/10.1007/s13164-015-0259-y

Chandler, M., Fritz, A. S., & Hala, S. (1989). Small-Scale Deceit: Deception as a Marker of Two-, Three-, and Four-Year-Olds' Early Theories of Mind. *Child Development*, *60*(6), 1263. https://doi.org/10.2307/1130919

Christensen, R. H. B. (2015). *ordinal—Regression Models for Ordinal Data*.

Cikara, M., Bruneau, E. G., & Saxe, R. R. (2011). Us and Them: Intergroup Failures of Empathy. *Current Directions in Psychological Science*, *20*(3), 149–153. https://doi.org/10.1177/0963721411408713

Corbit, J., McAuliffe, K., Callaghan, T. C., Blake, P. R., & Warneken, F. (2017). Children's collaboration induces fairness rather than generosity. *Cognition*, *168*, 344–356. https://doi.org/10.1016/j.cognition.2017.07.006

Cortes, B. P., Demoulin, S., Rodriguez, R. T., Rodriguez, A. P., & Leyens, J.-P. (2005). Infrahumanization or Familiarity? Attribution of Uniquely Human Emotions to the Self, the Ingroup, and the Outgroup. *Personality and Social Psychology Bulletin*, *31*(2), 243–253. https://doi.org/10.1177/0146167204271421

Dadds, M. R., Hunter, K., Hawes, D. J., Frost, A. D. J., Vassallo, S., Bunn, P., … Masry, Y. E. (2008). A measure of cognitive and affective empathy in children using parent ratings. *Child Psychiatry and Human Development*, *39*(2), 111–122. https://doi.org/10.1007/s10578-007-0075-4

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, *33*(1), 67–80. https://doi.org/10.1007/s00199-006-0153-z

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. *Nature*, *446*(7137), 794–796. https://doi.org/10.1038/nature05651

Decety, J., & Cacioppo, S. (2012). The speed of morality: A high-density electrical neuroimaging study. *Journal of Neurophysiology*, *108*(11), 3068–3072. https://doi.org/10.1152/jn.00473.2012

Decety, Jean, Jackson, P. L., Sommerville, J. A., Chaminade, T., & Meltzoff, A. N. (2004). The neural bases of cooperation and competition: An fMRI investigation. *NeuroImage*, *23*(2), 744–751. https://doi.org/10.1016/j.neuroimage.2004.05.025

Devine, R. T., & Hughes, C. (2016). Measuring theory of mind across middle childhood: Reliability and validity of the Silent Films and Strange Stories tasks. *Journal of Experimental Child Psychology*, *149*, 23–40. https://doi.org/10.1016/j.jecp.2015.07.011

Devine, R. T., White, N., Ensor, R., & Hughes, C. (2016). Theory of mind in middle childhood: Longitudinal associations with executive function and social competence. *Developmental Psychology*, *52*(5), 758–771. https://doi.org/10.1037/dev0000105

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *NeuroImage*, *55*(2), 705–712. https://doi.org/10.1016/j.neuroimage.2010.12.040

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, *25*(4), 1325–1335. https://doi.org/10.1016/j.neuroimage.2004.12.034

Engelmann, J. M., Herrmann, E., & Tomasello, M. (2012). Five-year olds, but not chimpanzees, attempt to manage their reputations. *PLoS ONE*, *7*(10), e48433. https://doi.org/10.1371/journal.pone.0048433

Engelmann, J. M., & Rapp, D. J. (2018). The influence of reputational concerns on children's prosociality. *Current Opinion in Psychology*, *20*, 92–95. https://doi.org/10.1016/j.copsyc.2017.08.024

Epley, N., Akalis, S., Waytz, A., & Cacioppo, J. T. (2008). Creating Social Connection Through Inferential Reproduction: Loneliness and Perceived Agency in Gadgets, Gods, and Greyhounds. *Psychological Science*, *19*(2), 114–120. https://doi.org/10.1111/j.1467-9280.2008.02056.x

Epley, N., & Waytz, A. (2010). Mind Perception. In *Handbook of Social Psychology*. John Wiley & Sons, Inc. Retrieved from http://dx.doi.org/10.1002/9780470561119.socpsy001014

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868.

FeldmanHall, O., Dalgleish, T., Thompson, R., Evans, D., Schweizer, S., & Mobbs, D. (2012). Differential neural circuitry and self-interest in real *vs* hypothetical moral decisions. *Social Cognitive and Affective Neuroscience*, *7*(7), 743–751. https://doi.org/10.1093/scan/nss069

FeldmanHall, O., Mobbs, D., Evans, D., Hiscox, L., Navrady, L., & Dalgleish, T. (2012). What we say and what we do: The relationship between real and hypothetical moral choices. *Cognition*, *123*(3), 434–441. https://doi.org/10.1016/j.cognition.2012.02.001

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly*, *50*(3), 274–290.

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition*, *57*(2), 109–128.

Flombaum, J. I., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology*, *15*(5), 447–452. https://doi.org/10.1016/j.cub.2004.12.076

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*(4), 708–724.

Galinsky, Adam D., Maddux, W. W., Gilin, D., & White, J. B. (2008). Why it pays to get inside the head of your opponent: The differential effects of perspective taking and empathy in negotiations. *Psychological Science*, *19*(4), 378–384. https://doi.org/10.1111/j.1467-9280.2008.02096.x

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of "theory of mind" in verbal and nonverbal tasks. *Neuropsychologia*, *38*(1), 11–21.

Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental Science*, *14*(5), 1012–1020. https://doi.org/10.1111/j.1467-7687.2011.01048.x

Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J., & Haxby, J. V. (2007). Two takes on the social brain: A comparison of theory of mind tasks. *Journal of Cognitive Neuroscience*, *19*(11), 1803–1814. https://doi.org/10.1162/jocn.2007.19.11.1803

Gonzalez, Antonya M., Steele, J. R., & Baron, A. S. (2017). Reducing children's implicit racial bias through exposure to positive out-group exemplars. *Child Development*, *88*(1), 123–130. https://doi.org/10.1111/cdev.12582

Gonzalez, Antonya Marie, Dunlop, W. L., & Baron, A. S. (2016). Malleability of implicit associations across development. *Developmental Science*. https://doi.org/10.1111/desc.12481

Greene, Joshua D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105–2108.

Greene, Joshua David. (2013). *Moral tribes: Emotion, reason, and the gap between us and them*. New York: The Penguin Press.

Hadwin, J., & Perner, J. (1991). Pleased and surprised: Children's cognitive theory of emotion. *British Journal of Developmental Psychology*, *9*(2), 215–234. https://doi.org/10.1111/j.2044-835X.1991.tb00872.x

Hala, S., Chandler, M., & Fritz, A. S. (1991). Fledgling theories of mind: Deception as a marker of three-year-olds' understanding of false belief. *Child Development*, *62*(1), 83–97. https://doi.org/10.1111/j.1467-8624.1991.tb01516.x

Hamann, K., Bender, J., & Tomasello, M. (2014). Meritocratic sharing is based on collaboration in 3-year-olds. *Developmental Psychology*, *50*(1), 121–128. https://doi.org/10.1037/a0032965

Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature*, *476*(7360), 328–331. https://doi.org/10.1038/nature10278

Happé, F. G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, *66*(3), 843–855.

Hare, B. (2001). Can competitive paradigms increase the validity of experiments on primate social cognition? *Animal Cognition*, *4*(3–4), 269–280. https://doi.org/10.1007/s100710100084

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*(4), 771–785. https://doi.org/10.1006/anbe.1999.1377

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, *61*(1), 139–151. https://doi.org/10.1006/anbe.2000.1518

Hare, B., Call, J., & Tomasello, M. (2006). Chimpanzees deceive a human competitor by hiding. *Cognition*, *101*(3), 495–514. https://doi.org/10.1016/j.cognition.2005.01.011

Hare, B., & Tomasello, M. (2004). Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. *Animal Behaviour*, *68*(3), 571–581. https://doi.org/10.1016/j.anbehav.2003.11.011

Harris, L. T., & Fiske, S. T. (2007). Social groups that elicit disgust are differentially processed in mPFC. *Social Cognitive and Affective Neuroscience*, *2*(1), 45–51. https://doi.org/10.1093/scan/nsl037

Harris, L. T., & Fiske, S. T. (2009). Social neuroscience evidence for dehumanised perception. *European Review of Social Psychology*, *20*(1), 192–231. https://doi.org/10.1080/10463280902954988

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, *10*(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4

Haslam, N., & Loughnan, S. (2014). Dehumanization and infrahumanization. *Annual Review of Psychology*, *65*(1), 399–423. https://doi.org/10.1146/annurev-psych-010213-115045

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, *57*(2), 243. https://doi.org/10.2307/1416950

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Heyes, C. M. (1998). Theory of mind in nonhuman primates. *The Behavioral and Brain Sciences*, *21*(1), 101–114; discussion 115-148.

Hrdy, S. B. (2009). *Mothers and others: The evolutionary origins of mutual understanding*. Cambridge, Mass: Belknap Press of Harvard University Press.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, *16*(2), 233–253. https://doi.org/10.1111/j.2044-835X.1998.tb00921.x

Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development*, *13*(4), 590–619.

Jackson, J. W. (1993). Realistic group conflict theory: A review and evaluation of the theoretical and empirical literature. *The Psychological Record*, *43*(3), 395–413.

Jenkins, A. C., & Mitchell, J. P. (2010). Mentalizing under uncertainty: Dissociated neural responses to ambiguous and unambiguous mental state inferences. *Cerebral Cortex*, *20*(2), 404–410. https://doi.org/10.1093/cercor/bhp109

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Kahneman, D. (2012). *Thinking, fast and slow*. London: Penguin Books.

Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition*, *109*(2), 224–234. https://doi.org/10.1016/j.cognition.2008.08.010

Kanngiesser, P., & Warneken, F. (2012). Young children consider merit when sharing resources with others. *PLoS ONE*, *7*(8), e43979. https://doi.org/10.1371/journal.pone.0043979

Karg, K., Schmelz, M., Call, J., & Tomasello, M. (2015). The goggles experiment: can chimpanzees use self-experience to infer what a competitor can see? *Animal Behaviour*, *105*, 211–221. https://doi.org/10.1016/j.anbehav.2015.04.028

Kelman, H. G. (1973). Violence without moral restraint: Reflections on the dehumanization of victims and victimizers. *Journal of Social Issues*, *29*(4), 25–61. https://doi.org/10.1111/j.1540-4560.1973.tb00102.x

Koster-Hale, J., Saxe, R., Dungan, J., & Young, L. L. (2013). Decoding moral judgments from neural representations of intentions. *Proceedings of the National Academy of Sciences*, *110*(14), 5648–5653. https://doi.org/10.1073/pnas.1207992110

Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology*, *90*(4), 543–555. https://doi.org/10.1037/0022-3514.90.4.543

Lee, K. (2013). Little Liars: Development of Verbal Deception in Children. *Child Development Perspectives*, *7*(2), 91–96. https://doi.org/10.1111/cdep.12023

Leimgruber, K. L., Shaw, A., Santos, L. R., & Olson, K. R. (2012). Young children are more generous when others are aware of their actions. *PLoS ONE*, *7*(10), e48292. https://doi.org/10.1371/journal.pone.0048292

Leyens, J.-P., Paladino, P. M., Rodriguez-Torres, R., Vaes, J., Demoulin, S., Rodriguez-Perez, A., & Gaunt, R. (2000). The emotional side of prejudice: The attribution of secondary emotions to ingroups and outgroups. *Personality and Social Psychology Review*, *4*(2), 186–197. https://doi.org/10.1207/S15327957PSPR0402_06

Lissek, S., Peters, S., Fuchs, N., Witthaus, H., Nicolas, V., Tegenthoff, M., … Brüne, M. (2008). Cooperation and deception recruit different subsets of the theory-of-mind network. *PLoS ONE*, *3*(4), e2023. https://doi.org/10.1371/journal.pone.0002023

Liu, B., Huang, Z., Xu, G., Jin, Y., Chen, Y., Li, X., … Jing, J. (2016). Altruistic sharing behavior in children: Role of theory of mind and inhibitory control. *Journal of Experimental Child Psychology*, *141*, 222–228. https://doi.org/10.1016/j.jecp.2015.09.010

Loewenstein, G. F., Thompson, L., & Bazerman, M. H. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social Psychology*, *57*(3), 426.

Lurz, R. W. (2011). *Mindreading animals: The debate over what animals know about other minds*. Cambridge, Mass: MIT Press.

Lyons, D. E., & Santos, L. R. (2006). Ecology, domain specificity, and the origins of theory of mind: Is competition the catalyst? *Philosophy Compass*, *1*(5), 481–492. https://doi.org/10.1111/j.1747-9991.2006.00032.x

MacLaren, R., & Olson, D. (1993). Trick or treat: Children’s understanding of surprise. *Cognitive Development*, *8*(1), 27–46. https://doi.org/10.1016/0885-2014(93)90003-N

Malle, B. F., & Holbrook, J. (2012). Is there a hierarchy of social inferences? The likelihood and speed of inferring intentionality, mind, and personality. *Journal of Personality and Social Psychology*, *102*(4), 661–684. https://doi.org/10.1037/a0026790

McAuliffe, K., Blake, P. R., Kim, G., Wrangham, R. W., & Warneken, F. (2013). Social influences on inequity aversion in children. *PLoS ONE*, *8*(12), e80966. https://doi.org/10.1371/journal.pone.0080966

McAuliffe, K., Blake, P. R., Steinbeis, N., & Warneken, F. (2017). The developmental foundations of human fairness. *Nature Human Behaviour*, *1*, 0042.

McAuliffe, K., & Dunham, Y. (2016). Group bias in cooperative norm enforcement. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150073. https://doi.org/10.1098/rstb.2015.0073

McPherson Frantz, C., & Janoff-Bulman, R. (2000). Considering Both Sides: The Limits of Perspective Taking. *Basic and Applied Social Psychology*, *22*(1), 31–42. https://doi.org/10.1207/S15324834BASP2201_4

Melis, A. P., Call, J., & Tomasello, M. (2006). Chimpanzees (Pan troglodytes) conceal visual and auditory information from others. *Journal of Comparative Psychology*, *120*(2), 154–162. https://doi.org/10.1037/0735-7036.120.2.154

Morewedge, C. K. (2009). Negativity bias in attribution of external agency. *Journal of Experimental Psychology: General*, *138*(4), 535–545. https://doi.org/10.1037/a0016796

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science (New York, N.Y.)*, *308*(5719), 255–258. https://doi.org/10.1126/science.1107621

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. *Frontiers in Neuroinformatics*, *10*. https://doi.org/10.3389/fninf.2016.00027

Opotow, S. (1990). Moral exclusion and injustice: An introduction. *Journal of Social Issues*, *46*(1), 1–20. https://doi.org/10.1111/j.1540-4560.1990.tb00268.x

Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology*, *1*(1), 33–60.

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental Psychology*, *43*(4), 974–990. https://doi.org/10.1037/0012-1649.43.4.974

Perner, J, & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences*, *3*(9), 337–344.

Perner, Josef, & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences*, *16*(10), 519–525. https://doi.org/10.1016/j.tics.2012.08.004

Peterson, C. C., Slaughter, V., Peterson, J., & Premack, D. (2013). Children with autism can track others' beliefs in a competitive game. *Developmental Science*, *16*(3), 443–450. https://doi.org/10.1111/desc.12040

Phillips, W., Barnes, J. L., Mahajan, N., Yamaguchi, M., & Santos, L. R. (2009). 'Unwilling' versus 'unable': capuchin monkeys' (*Cebus apella*) understanding of human intentional action: Unwilling vs. unable in capuchin monkeys. *Developmental Science*, *12*(6), 938–945. https://doi.org/10.1111/j.1467-7687.2009.00840.x

Povinelli, D. J., Rulf, A. B., & Bierschwale, D. T. (1994). Absence of knowledge attribution and self-recognition in young chimpanzees (Pan troglodytes).

Journal of Comparative Psychology, *108*(1), 74–80. https://doi.org/10.1037/0735-7036.108.1.74

Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: suspiciously human? *Trends in Cognitive Sciences*, *7*(4), 157–160. https://doi.org/10.1016/S1364-6613(03)00053-6

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(04), 515. https://doi.org/10.1017/S0140525X00076512

Raihani, N. J., & McAuliffe, K. (2012). Human punishment is motivated by inequity aversion, not a desire for reciprocity. *Biology Letters*, *8*(5), 802–804. https://doi.org/10.1098/rsbl.2012.0470

Reis, H. T., & Gruzen, J. (1976). On mediating equity, equality, and self-interest: The role of self-presentation in social exchange. *Journal of Experimental Social Psychology*, *12*(5), 487–503. https://doi.org/10.1016/0022-1031(76)90079-2

Rizzo, M. T., & Killen, M. (2016). Children's understanding of equity in the context of inequality. *British Journal of Developmental Psychology*, *34*(4), 569–581. https://doi.org/10.1111/bjdp.12150

Rochat, P., Dias, M. D. G., Guo Liping, Broesch, T., Passos-Ferreira, C., Winning, A., & Berg, B. (2009). Fairness in distributive justice by 3- and 5-year-olds across seven cultures. *Journal of Cross-Cultural Psychology*, *40*(3), 416–442. https://doi.org/10.1177/0022022109332844

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and U.S. preschoolers. *Psychological Science*, *17*(1), 74–81. https://doi.org/10.1111/j.1467-9280.2005.01667.x

Santiesteban, I., Banissy, M. J., Catmur, C., & Bird, G. (2015). Functional lateralization of temporoparietal junction - imitation inhibition, visual perspective-taking and theory of mind. *European Journal of Neuroscience*, *42*(8), 2527–2533. https://doi.org/10.1111/ejn.13036

Santos, L. R., Nissen, A. G., & Ferrugia, J. A. (2006). Rhesus monkeys, Macaca mulatta, know what others can and cannot hear. *Animal Behaviour*, *71*(5), 1175–1181. https://doi.org/10.1016/j.anbehav.2005.10.007

Saxe, R, & Kanwisher, N. (2003). People thinking about thinking peopleThe role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*(4), 1835–1842. https://doi.org/10.1016/S1053-8119(03)00230-1

Saxe, R, & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, *43*(10), 1391–1399. https://doi.org/10.1016/j.neuropsychologia.2005.02.013

Saxe, Rebecca, Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, *55*, 87–124. https://doi.org/10.1146/annurev.psych.55.090902.142044

Saxe, Rebecca, & Powell, L. J. (2006). It's the thought that counts: Specific brain regions for one component of theory of mind. *Psychological Science*, *17*(8), 692–699. https://doi.org/10.1111/j.1467-9280.2006.01768.x

Schmelz, M., Call, J., & Tomasello, M. (2011). Chimpanzees know that others make inferences. *Proceedings of the National Academy of Sciences*, *108*(7), 3077–3079. https://doi.org/10.1073/pnas.1000469108

Schmelz, Martin, & Call, J. (2016). The psychology of primate cooperation and competition: a call for realigning research agendas. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150067. https://doi.org/10.1098/rstb.2015.0067

Schmidt, M. F. H., & Sommerville, J. A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. *PLoS ONE*, *6*(10), e23223. https://doi.org/10.1371/journal.pone.0023223

Schug, J., Takagishi, H., Benech, C., & Okada, H. (2016). The development of theory of mind and positive and negative reciprocity in preschool children. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00888

Schurz, M., Aichhorn, M., Martin, A., & Perner, J. (2013). Common brain areas engaged in false belief reasoning and visual perspective taking: a meta-analysis of functional brain imaging studies. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00712

Scott, R. M., & Baillargeon, R. (2017). Early False-Belief Understanding. *Trends in Cognitive Sciences*, *21*(4), 237–249. https://doi.org/10.1016/j.tics.2017.01.012

Shaw, A., Montinari, N., Piovesan, M., Olson, K. R., Gino, F., & Norton, M. I. (2014). Children develop a veil of fairness. *Journal of Experimental Psychology: General*, *143*(1), 363–375. https://doi.org/10.1037/a0031247

Sheskin, M., Bloom, P., & Wynn, K. (2014). Anti-equality: Social comparison in young children. *Cognition*, *130*(2), 152–156. https://doi.org/10.1016/j.cognition.2013.10.008

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience & Biobehavioral Reviews*, *30*(6), 855–863. https://doi.org/10.1016/j.neubiorev.2006.06.011

Singer, T. (2009). Understanding others: Brain mechanisms of theory of mind and empathy. *Neuroeconomics: Decision Making and the Brain*, 251–268.

Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness? *Psychological Science*, *23*(2), 196–204. https://doi.org/10.1177/0956797611422072

Smith, C. E., Blake, P. R., & Harris, P. L. (2013). I should but I won't: Why young children endorse norms of fair sharing but do not follow them. *PloS One*, *8*(3), e59510. https://doi.org/10.1371/journal.pone.0059510

Struch, N., & Schwartz, S. H. (1989). Intergroup aggression: Its predictors and distinctness from in-group bias. *Journal of Personality and Social Psychology*, *56*(3), 364–373.

Takagishi, H., Koizumi, M., Fujii, T., Schug, J., Kameshima, S., & Yamagishi, T. (2014). The role of cognitive and emotional perspective taking in economic decision making in the Ultimatum Game. *PLoS ONE*, *9*(9), e108462. https://doi.org/10.1371/journal.pone.0108462

Thornton, A., & McAuliffe, K. (2015). Cognitive consequences of cooperative breeding? A critical appraisal: Cognitive consequences of cooperative breeding. *Journal of Zoology*, *295*(1), 12–22. https://doi.org/10.1111/jzo.12198

Todd, A. R., & Galinsky, A. D. (2014). Perspective-Taking as a Strategy for Improving Intergroup Relations: Evidence, Mechanisms, and Qualifications: Perspective-Taking and Intergroup Relations. *Social and Personality Psychology Compass*, *8*(7), 374–387. https://doi.org/10.1111/spc3.12116

Tsoi, L., Dungan, J. A., Chakroff, A., & Young, L. L. (2018). Neural substrates for moral judgments of psychological versus physical harm. *Social Cognitive and Affective Neuroscience*. https://doi.org/10.1093/scan/nsy029

Tsoi, L., Dungan, J., Waytz, A., & Young, L. (2016). Distinct neural patterns of social cognition for cooperation versus competition. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2016.04.069

Tsoi, L., Hamlin, J. K., Baron, A. S., Waytz, A., & Young, L. (under review). Developmental evidence for enhanced theory of mind for competitive versus cooperative contexts.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*(3), 383–403. https://doi.org/10.1037/0033-2909.134.3.383

Valdesolo, P., & DeSteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion*, *11*(2), 262–266. https://doi.org/10.1037/a0021302

Wagner, D. D., Kelley, W. M., & Heatherton, T. F. (2011). Individual differences in the spontaneous recruitment of brain regions supporting mental state understanding when viewing natural social scenes. *Cerebral Cortex*, *21*(12), 2788–2796. https://doi.org/10.1093/cercor/bhr074

Waytz, A., Epley, N., & Cacioppo, J. T. (2010). Social Cognition Unbound: Insights Into Anthropomorphism and Dehumanization. *Current Directions in Psychological Science*, *19*(1), 58–62. https://doi.org/10.1177/0963721409359302

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, *14*(8), 383–388. https://doi.org/10.1016/j.tics.2010.05.006

Waytz, A., Morewedge, C. K., Epley, N., Monteleone, G., Gao, J.-H., & Cacioppo, J. T. (2010a). Making sense by making sentient: Effectance motivation increases anthropomorphism. *Journal of Personality and Social Psychology*, *99*(3), 410–435. https://doi.org/10.1037/a0020240

Waytz, A., Morewedge, C. K., Epley, N., Monteleone, G., Gao, J.-H., & Cacioppo, J. T. (2010b). Making sense by making sentient: Effectance motivation increases anthropomorphism. *Journal of Personality and Social Psychology*, *99*(3), 410–435. https://doi.org/10.1037/a0020240

Waytz, A., & Young, L. (2014). Two motivations for two dimensions of mind. *Journal of Experimental Social Psychology*, *55*, 278–283. https://doi.org/10.1016/j.jesp.2014.08.001

Wellman, H. M., & Banerjee, M. (1991). Mind and emotion: Children's understanding of the emotional consequences of beliefs and desires. *British Journal of Developmental Psychology*, *9*(2), 191–214. https://doi.org/10.1111/j.2044-835X.1991.tb00871.x

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655–684.

Wheatley, T., Milleville, S. C., & Martin, A. (2007). Understanding animate agents: Distinct roles for the social network and mirror system. *Psychological Science*, *18*(6), 469–474. https://doi.org/10.1111/j.1467-9280.2007.01923.x

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*(5), 297–333. https://doi.org/10.1037/h0040934

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and Cooperation. *Psychological Science*, *20*(1), 1–5. https://doi.org/10.1111/j.1467-9280.2008.02253.x

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128.

Woodruff, G., & Premack, D. (1979). Intentional communication in the chimpanzee: The development of deception. *Cognition*, *7*(4), 333–362. https://doi.org/10.1016/0010-0277(79)90021-0

Yirmiya, N., Erel, O., Shaked, M., & Solomonica-Levi, D. (1998). Meta-analyses comparing theory of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. *Psychological Bulletin*, *124*(3), 283–307.

Yoder, K. J., & Decety, J. (2014). The Good, the Bad, and the Just: Justice sensitivity predicts neural response during moral evaluation of actions performed by others. *Journal of Neuroscience*, *34*(12), 4161–4166. https://doi.org/10.1523/JNEUROSCI.4648-13.2014

Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences*, *104*(20), 8235–8240. https://doi.org/10.1073/pnas.0701408104

Young, Liane, & Saxe, R. (2009a). An fMRI investigation of spontaneous mental state inference for moral judgment. *Journal of Cognitive Neuroscience*, *21*(7), 1396–1405. https://doi.org/10.1162/jocn.2009.21137

Young, Liane, & Saxe, R. (2009b). Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia*, *47*(10), 2065–2072. https://doi.org/10.1016/j.neuropsychologia.2009.03.020

Zaki, J., & Ochsner, K. (2016). Empathy. In *Handbook of Emotion* (4th ed., pp. 871–884). New York: Guilford.

4.7 APPENDIX

4.7.1 Script for the experimental task

Key:
A = actor
R = recipient
E = experimenter

THEORY OF MIND CONDITION

Before the kids sit down

Hi, my name is _______, what's your name? Great! Before we get started with the game, there's something I want to show you. This is where the game will take place. If you come over to this side, you know what's on this side, and you know what's on the other side. Now come around to this side. If you are on this side, you know what's on this side, and you know what's on the other side. Does this make sense? Great. So you're going to sit on this cushion *(assign actor)*, and you're going to sit on this cushion *(assign recipient)* and now I'll explain the game.

Intro and preliminary checks

Well, [names], today we're going to be playing a game with Skittles and at the end of the game, you'll get to take home some Skittles. So, I have a very important question for you: do you like Skittles? Great! I think you're going to like this game.

Before we get started, though, I want to make sure you know that you don't have to play this game if you don't want to, and if at any point you want to stop playing the game, just let me know and we'll stop, and you'll still get to take home some Skittles. Do you want to play? Great!

Demonstration Phase

In this game, you'll [A] be in charge of making *ALL* the decisions and you'll be making decisions using this toy. So let me show you how the toy works. I'm going to put some Skittles down on these trays and the ones that I put here are for [recipient] and the ones that I put here are for [actor]. So I'll put some on R's tray and some on A's tray.

Now, [actor's name], you can either pull the green handle or the red handle.

[order of presentation of handles counterbalanced; this is noted on the data sheets]

(Green handle): If you pull the green handle, look what happens *(E pulls handle here)*. The Skittles fall into your bowls and you get to keep those. I'm giving each of you a paper bag so you can put your candies from these bowls into your paper bag. Here's a bag for you *(give bag to recipient)* and here's a bag for you *(give bag to actor)*. Let's move your candies to your paper bag now!

(Red handle) So let's see what the red handle does. I'll put some candy down on this side and some on this side *(E places 1 on each side again)*. If you pull the red handle, look what happens *(E pulls handle here)*. They drop into the middle and nobody gets those Skittles.

(E places 1 candy on each side to demonstrate handles; candy on recipient's side first. We always put candy on the recipient's tray first to draw the child's attention to that side – otherwise they could just focus on their own reward.)

In a moment we'll do some practice to get used to the game. Before we do, though, there are three important rules of the game: First, when this stick is on the trays, please don't touch the toy. When I lift the stick, you [actor] can pull one of the handles. Ok? Second, there is no talking in this game. Third, please don't eat any candy during the game. Instead, save the treats until after the game. After the game you can take them home. Ok?

Practice Trials

(There are 3 practice trials where children must pull the handles themselves to experience how each one works. No feedback is given for these trials. Trial 1 is always 1 candy each. The next two trials must be counterbalanced – one is an allocation of 1 candy to the actor and zero to the recipient; the other is zero for the actor and 1 for the recipient. The order of these second two trials is noted on the data sheets.)

Okay, let's do some practice to make sure you understand how this works.

(Put stick across the trays)

Alright, now the stick is on the trays so no one should be touching the toy. Good. Now, I'm going to put one on [recipient's] side and one on your side.

Now I lift the stick [Lift the stick], and [actor], which handle do you want to pull?

(If child does not pull a handle within about 5 seconds tell them – "remember, the green handle makes the candies fall into the two bowls and the red handle makes the candies fall into the middle bowl [try to counterbalance in line with demonstration order]).

(State the outcome)
OK, you each get one candy OR

OK, no one gets any candy this time.
(When they accept the offer, make sure they move the candy to their paper bags.)

(Reset trays and place stick across)
Now let's try another one.

(The next two trials will be counterbalanced – refer to the data sheets.)
I'm going to put one on [recipient's] side and none on your side. [Lift the stick] Now, which handle do you want to pull?
(State the outcome)
OK, R gets one candy and A does not get any candy OR
OK, no one gets any candy this time

I'm going to put none on [recipient's] side and one on your side. [Lift the stick] Now, which handle do you want to pull?
(State the outcome)
OK, R does not get any candy and A gets one candy OR
OK, no one gets any candy this time

(If the actor only pulls green or red for all 3 trials, add an additional trial of 1-1 and say, "Now can you show me which handle to pull to make the candy drop into the middle" if they have not pulled red (for example) or " . . . to make the candy drop into the outside bowls" if they have not pulled green. This is to ensure that the child has experienced the result of pulling each handle.)

(After the 3 (or more if needed) practice trials, put down a 1-1 allocation for the comp questions)

Before we make a decision, I want to ask the both of you some questions. R, you know how much candy is on your side, but do you know how much candy is on A's side? [R should say yes or nod]. That's right, and does A know how much candy is on your side? Exactly. You know how much is on A's side and A knows how much is on your side.

And A, you know how much candy is on your side, but do you know how much candy is on R's side? [A should say yes or nod]. That's right. And does R know how much candy is on your side? Exactly. You know ho much is on R's side and R knows how much is on your side.

For fun, I'm going to pull the green handle, and you each get one candy.

Are you ready to play the game?

Test Trials

Each pair of kids only sees one type of unequal trial in the session – either disadvantageous inequity (1 for actor, 4 for recipient) or advantageous inequity (4 for

actor, 1 for recipient). This is really important. If the direction of inequity is accidentally reversed, the session should be considered invalid.
For each pair, there will be 12 trials. 6 equal trials (1-1) and 6 unequal (4-1 or 1-4). The sequence of the equal and unequal trials is randomized beforehand and written on the data sheets so that the experimenter can just follow what is pre-written. For each trial, E states the number of candies placed on each side and lifts the stick. If the child doesn't pull after the stick is lifted after about 5 seconds, put the stick down again and say, "which handle do you want to pull" then lift the stick. If the actor rejects the offer, E says "No one gets any [treat] that time." If the actor accepts the offer, E states how many each player received, i.e., "So [recipient] gets 4 and [actor] get 1" Or "[recipient] gets 1 and [actor] gets 1.
If necessary, remind children that they are not allowed to eat the candy during the game and are discouraged from talking during the game.
After the test trials are completed, E says:

OK, that's the end of the decision-making part of this study! A, I have some questions for you now, and then afterwards, you will be doing another task with [other experimenter] (point at other experimenter). R, you are now done with the study, and you can take your Skittles home! Thank you for playing this game with us!

Post-Game questions

E asks justification questions of the actor for whatever decisions were made on the unequal trials. Younger children often fail to respond or say I don't know. Just record 'no response.' We are interested in the age at which children start to give explicit justifications for their actions.

Why red when unequal?
Why green when unequal?

CONTROL CONDITION

Before the kids sit down

Hi, my name is _______, what's your name? Great! Before we get started with the game, there's something I want to show you. This is where the game will take place. If you come around to this side, you can see that there is a (color of bowl) bowl on this side. Now come around to this side. If you are on this side, you can see that there's a (color of bowl) bowl on this side. Does this make sense? Great. So you're going to sit on this cushion (assign actor), and you're going to sit on this cushion (assign recipient) and now I'll explain the game.

Intro and preliminary checks

Well, [names], today we're going to be playing a game with Skittles and at the end of the game, you'll get to take home some Skittles. So, I have a very important question for you: do you like Skittles? Great! I think you're going to like this game.

Before we get started, though, I want to make sure you know that you don't have to play this game if you don't want to, and if at any point you want to stop playing the game, just let me know and we'll stop, and you'll still get to take home some Skittles. Do you want to play? Great!

Demonstration Phase

In this game, you'll [A] be in charge of making *ALL* the decisions and you'll be making decisions using this toy. So let me show you how the toy works. I'm going to put some Skittles down on these trays and the ones that I put here are for [recipient] and the ones that I put here are for [actor]. So I'll put some on R's tray and some on A's tray.

Now, [actor's name], you can either pull the green handle or the red handle.

[order of presentation of handles counterbalanced; this is noted on the data sheets]

(Green handle): If you pull the green handle, look what happens *(E pulls handle here)*. The Skittles fall into your bowls and you get to keep those. I'm giving each of you a paper bag so you can put your candies from these bowls into your paper bag. Here's a bag for you *(give bag to recipient)* and here's a bag for you *(give bag to actor)*. Why don't you move your candies to your paper bag now?

(Red handle) So lets see what the red handle does. I'll put some candy down on this side and some on this side *(E places 1 on each side again)*. If you pull the red handle look what happens *(E pulls handle here)*. They drop into the middle and nobody gets those Skittles.

(E places 1 candy on each side to demonstrate handles; candy on recipient's side first.

We always put candy on the recipient's tray first to draw the child's attention to that side – otherwise they could just focus on their own reward.)

In a moment we'll do some practice to get used to the game. Before we do, though, there are three important rules of the game: First, when this stick is on the trays, please don't touch the toy. When I lift the stick, you [actor] can pull one of the handles. Ok? Second, there is no talking in this game. Third, please don't eat any candy during the game. Instead, save the treats until after the game. After the game you can take them home. Ok?

Practice Trials

(There are 3 practice trials where children must pull the handles themselves to experience how each one works. No feedback is given for these trials. Trial 1 is always 1 candy each. The next two trials must be counterbalanced – one is an allocation of 1 candy to the actor and zero to the recipient; the other is zero for the actor and 1 for the recipient. The order of these second two trials is noted on the data sheets.)

Okay, let's do some practice to make sure you understand how this works.

(Put stick across the trays)

Alright, now the stick is on the trays so no one should be touching the toy. Good. Now, I'm going to put one on [recipient's] side and one on your side.

Now I lift the stick [Lift the stick], and [actor], which handle do you want to pull?

(If child does not pull a handle within about 5 seconds tell them – "remember, the green handle makes the candies fall into the two bowls and the red handle makes the candies fall into the middle bowl [try to counterbalance in line with demonstration order]).

(State the outcome)
OK, you each get one candy OR
OK, no one gets any candy this time.
(When they accept the offer, make sure they move the candy to their paper bags.)

(Reset trays and place stick across)
Now let's try another one.

(The next two trials will be counterbalanced – refer to the data sheets.)
I'm going to put one on [recipient's] side and none on your side. [Lift the stick] Now, which handle do you want to pull?
(State the outcome)
OK, R gets one candy and A does not get any candy OR
OK, no one gets any candy this time

I'm going to put none on [recipient's] side and one on your side. [Lift the stick] Now, which handle do you want to pull?
(State the outcome)
OK, R does not get any candy and A gets one candy OR
OK, no one gets any candy this time

(If the actor only pulls green or red for all 3 trials, add an additional trial of 1-1 and say, "Now can you show me which handle to pull to make the candy drop into the middle" if they have not pulled red (for example) or " . . . to make the candy drop into the outside bowls" if they have not pulled green. This is to ensure that the child has experienced the result of pulling each handle.)

(After the 3 (or more if needed) practice trials, put down a 1-1 allocation for the comp questions)

Before you make a decision, I want to ask the both of you a question. R, you have an orange bowl. What color is A's bowl? That's right, A has a X bowl. And A, you have a XX bowl what color is R's bowl. That's right R has a XX bowl.

For fun, I'm going to pull the green handle, and you each get one candy.

Are you ready to play the game?

Test Trials

Each pair of kids only sees one type of unequal trial in the session – either disadvantageous inequity (1 for actor, 4 for recipient) or advantageous inequity (4 for actor, 1 for recipient). This is really important. If the direction of inequity is accidentally reversed, the session should be considered invalid.
For each pair, there will be 12 trials. 6 equal trials (1-1) and 6 unequal (4-1 or 1-4). The sequence of the equal and unequal trials is randomized beforehand and written on the data sheets so that the experimenter can just follow what is pre-written. For each trial, E states the number of candies placed on each side and lifts the stick. If the child doesn't pull after the stick is lifted after about 5 seconds, put the stick down again and say, "which handle do you want to pull" then lift the stick. If the actor rejects the offer, E says "No one gets any [treat] that time." If the actor accepts the offer, E states how many each player received, i.e., "So [recipient] gets 4 and [actor] get 1" Or "[recipient] gets 1 and [actor] gets 1.
If necessary, remind children that they are not allowed to eat the candy during the game and are discouraged from talking during the game.
After the test trials are completed, E says:

OK, that's the end of the decision-making part of this study! A, I have some questions for you now, and then afterwards, you will be doing another task with [other experimenter] (point at other experimenter). R, you are now done with the study, and you can take your Skittles home! Thank you for playing this game with us!

Post-Game questions

E asks justification questions of the actor for whatever decisions were made on the unequal trials. Younger children often fail to respond or say I don't know. Just record 'no response.' We are interested in the age at which children start to give explicit justifications for their actions.

Why red when unequal?
Why green when unequal?

4.7.2 Reading the Mind in the Eyes Test (child version)

The task was conducted on an iPad. The 28 items were presented in a randomized order.

Sources:

Child version: Baron-Cohen, S., Wheelwright, S., Spong, A., Scahill, V., & Lawson, J. (2001). Are intuitive physics and intuitive psychology independent? A test with children with Asperger Syndrome. *Journal of Developmental and Learning Disorders*, *5*(1), 47–78.

Adult version: Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*(2), 241–251.

Variations in other languages: https://www.autismresearchcentre.com/arc_tests/

Instructions:

In this folder I've got lots of pictures of people's eyes. Each picture has four words round it. I want you to look carefully at the picture and then choose the word that best describes what the person in the picture is thinking or feeling. Let's have a go with this one (practice item). Look at this person. Do you think he is feeling jealous, scared, relaxed or hate (point to words as they are read)? Make sure child picks one of the options and give encouraging feedback without revealing whether they are right or wrong.

OK, let's have a go at the rest of them. You might find some of them quite easy and some of them quite hard, so don't worry if it's not always easy to choose the best word. I'll read all the words for you so you don't need to worry about that. If you really can't choose the best word, you can have a guess. Proceed with the test items in exactly the same way as the practice item.

practice

jealous scared

relaxed hate

1

kind cross

surprised hate

sad

cross

surprised

unkind

worried sad

surprised friendly

4

upset excited

surprised relaxed

5

relaxed joking

making somebody do something feeling sorry

6

hate bored

unkind worried

joking feeling sorry

interested bored

happy friendly

remembering angry

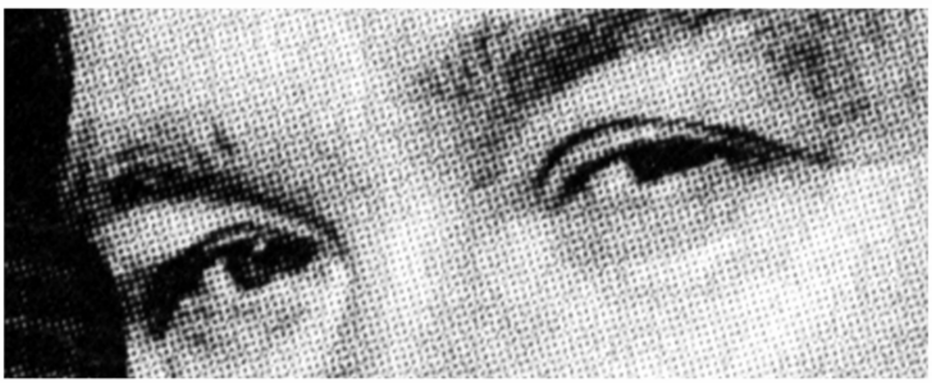
9

thinking about something annoyed

surprised hate

10

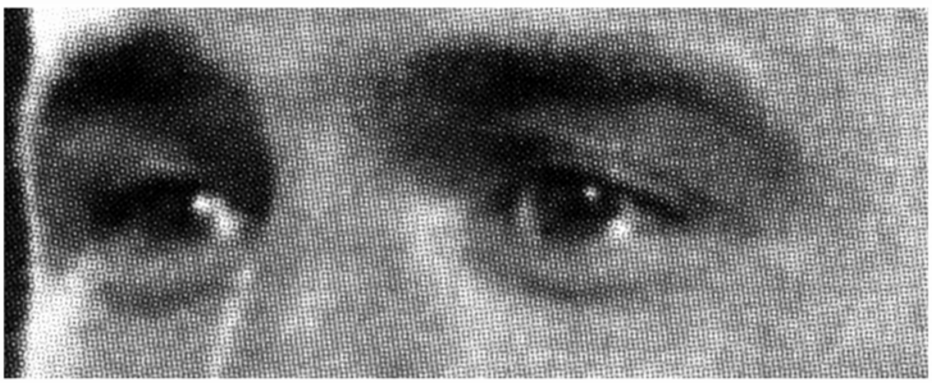
not believing sad



shy kind

11

hoping disgusted

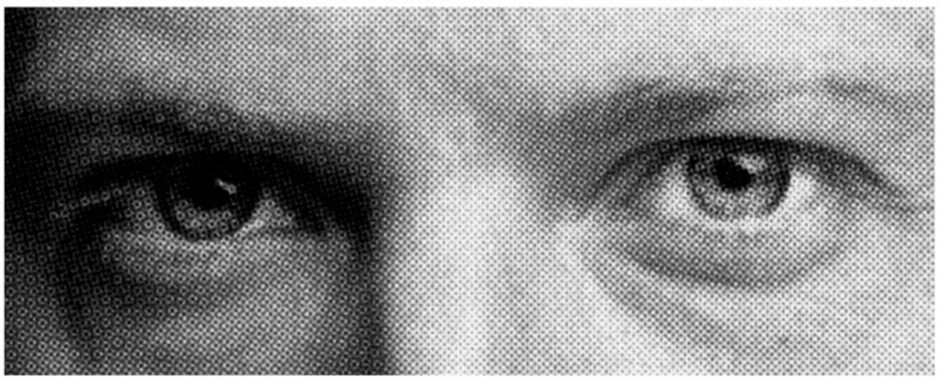


angry bossy

12

joking

confused



sad

serious

upset happy

excited thinking about something

14

happy excited

kind thinking about something

not believing wanting to play

friendly relaxed

16

made up her mind

bored

surprised

joking

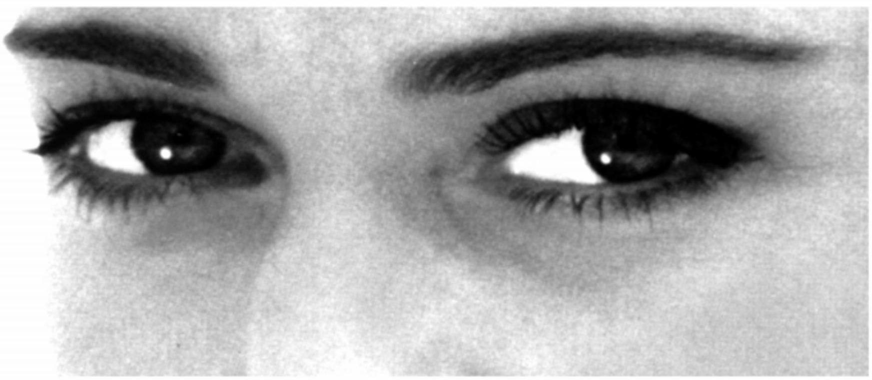
a bit worried angry

unkind friendly

friendly bossy

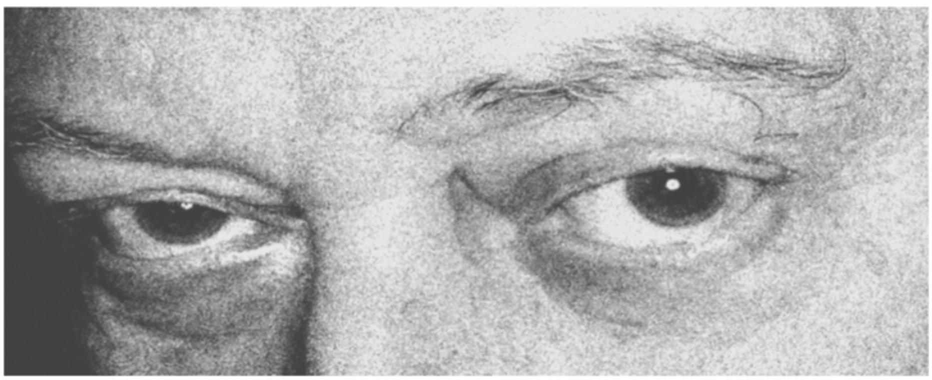
thinking about something sad angry

daydreaming sad



angry interested

excited surprise



kind not pleased

joking relaxed

happy interested

thinking about something playful

surprised kind

sure about something joking

surprised happy

ashamed serious

confused surprised

guilty worried

daydreaming shy

sorry relaxed

joking nervous

excited pleased

ashamed not believing

happy bored

hate disgust

4.7.3 Griffith Empathy Measure

The questionnaire was printed out on paper. We had 2 sets, each with questions presented in a different order. The sets (A or B) were presented in a counterbalanced manner across participants.

Parents were asked to fill out the questionnaire to the best of their abilities (and they could opt not to fill it out).

Source:

Dadds, M. R., Hunter, K., Hawes, D. J., Frost, A. D. J., Vassallo, S., Bunn, P., … Masry, Y. E. (2008). A measure of cognitive and affective empathy in children using parent ratings. *Child Psychiatry and Human Development*, *39*(2), 111–122. https://doi.org/10.1007/s10578-007-0075-4

SET A

Questionnaire for Parents

Please circle the answer that best describes your agreement with each statement.

This is how the scale will look like:

strongly disagree								strongly agree
-4	-3	-2	-1	0	1	2	3	4

My child gets sad when watching sad movies or TV.

-4 -3 -2 -1 0 1 2 3 4

My child can't understand why other people get upset.

-4 -3 -2 -1 0 1 2 3 4

My child gets upset seeing another child being punished for being naughty.

-4 -3 -2 -1 0 1 2 3 4

My child becomes sad when other children are sad.

-4 -3 -2 -1 0 1 2 3 4

My child doesn't understand why other people cry out of happiness.

-4 -3 -2 -1 0 1 2 3 4

My child gets upset when seeing another child being hurt.

-4 -3 -2 -1 0 1 2 3 4

My child becomes nervous when other children around them are nervous.

-4 -3 -2 -1 0 1 2 3 4

Again, this is how the scale looks like:

strongly disagree								strongly agree
-4	-3	-2	-1	0	1	2	3	4

My child rarely understands why other people cry.

-4 -3 -2 -1 0 1 2 3 4

My child laughs when seeing another child laugh.

-4 -3 -2 -1 0 1 2 3 4

My child acts happy when another person is acting happy.

-4 -3 -2 -1 0 1 2 3 4

My child likes to watch people open presents, even if not one for him/her.

-4 -3 -2 -1 0 1 2 3 4

My child treats cats and dogs like they have feelings.

-4 -3 -2 -1 0 1 2 3 4

My child gets upset when another person is acting upset.

-4 -3 -2 -1 0 1 2 3 4

My child cries or gets upset when seeing another child cry.

-4 -3 -2 -1 0 1 2 3 4

My child doesn't seem to notice when I get sad.

-4 -3 -2 -1 0 1 2 3 4

My child feels sorry for another child who is upset.

-4 -3 -2 -1 0 1 2 3 4

Again, this is how the scale looks like:

strongly disagree									strongly agree
	-4	-3	-2	-1	0	1	2	3	4

My child would eat the last cookie, even when they know someone else wants it.

-4 -3 -2 -1 0 1 2 3 4

My child gets upset when seeing an animal being hurt.

-4 -3 -2 -1 0 1 2 3 4

My child reacts badly when they see people kiss and hug in public.

-4 -3 -2 -1 0 1 2 3 4

My child can continue to feel okay even if people around are upset.

-4 -3 -2 -1 0 1 2 3 4

My child gets sad to see a child with no one to play with.

-4 -3 -2 -1 0 1 2 3 4

My child feels sad for people who are physically disabled.

-4 -3 -2 -1 0 1 2 3 4

My child seems to react to the moods of people around them.

-4 -3 -2 -1 0 1 2 3 4

SET B

Questionnaire for Parents

Please circle the answer that best describes your agreement with each statement.

This is how the scale will look like:

strongly disagree								strongly agree
-4	-3	-2	-1	0	1	2	3	4

My child gets upset when seeing an animal being hurt.

-4 -3 -2 -1 0 1 2 3 4

My child feels sad for people who are physically disabled.

-4 -3 -2 -1 0 1 2 3 4

My child doesn't seem to notice when I get sad.

-4 -3 -2 -1 0 1 2 3 4

My child likes to watch people open presents, even if not one for him/her.

-4 -3 -2 -1 0 1 2 3 4

My child becomes nervous when other children around them are nervous.

-4 -3 -2 -1 0 1 2 3 4

My child would eat the last cookie, even when they know someone else wants it.

-4 -3 -2 -1 0 1 2 3 4

My child doesn't understand why other people cry out of happiness.

-4 -3 -2 -1 0 1 2 3 4

Again, this is how the scale looks like:

strongly disagree									strongly agree
	-4	-3	-2	-1	0	1	2	3	4

My child gets upset when another person is acting upset.

-4 -3 -2 -1 0 1 2 3 4

My child laughs when seeing another child laugh.

-4 -3 -2 -1 0 1 2 3 4

My child treats cats and dogs like they have feelings.

-4 -3 -2 -1 0 1 2 3 4

My child gets sad to see a child with no one to play with.

-4 -3 -2 -1 0 1 2 3 4

My child acts happy when another person is acting happy.

-4 -3 -2 -1 0 1 2 3 4

My child reacts badly when they see people kiss and hug in public.

-4 -3 -2 -1 0 1 2 3 4

My child becomes sad when other children are sad.

-4 -3 -2 -1 0 1 2 3 4

My child gets upset seeing another child being punished for being naughty.

-4 -3 -2 -1 0 1 2 3 4

My child seems to react to the moods of people around them.

-4 -3 -2 -1 0 1 2 3 4

Again, this is how the scale looks like:

strongly disagree									strongly agree
	-4	-3	-2	-1	0	1	2	3	4

My child feels sorry for another child who is upset.

-4 -3 -2 -1 0 1 2 3 4

My child can't understand why other people get upset.

-4 -3 -2 -1 0 1 2 3 4

My child gets upset when seeing another child being hurt.

-4 -3 -2 -1 0 1 2 3 4

My child cries or gets upset when seeing another child cry.

-4 -3 -2 -1 0 1 2 3 4

My child gets sad when watching sad movies or TV.

-4 -3 -2 -1 0 1 2 3 4

My child can continue to feel okay even if people around are upset.

-4 -3 -2 -1 0 1 2 3 4

My child rarely understands why other people cry.

-4 -3 -2 -1 0 1 2 3 4

4.8 SUPPLEMENTARY MATERIAL

4.8.1 Participants

Table S1. Final sample breakdown (N=123) by Inequity Type, Decider Age Group, Decider Gender, and Condition.

			Control	Theory of Mind
Advantageous Inequity (AI)	6&7	Female	5	11
		Male	10	6
	8&9	Female	8	7
		Male	4	11
Disadvantageous Inequity (DI)	6&7	Female	7	8
		Male	7	8
	8&9	Female	7	8
		Male	8	8

4.8.2 Full model

Table S2. Model output from the full model. Estimates and standard errors (s.e.) of fixed effects in Generalized Linear Mixed Models predicting children's rejection behaviors. Baselines were set as follows: Distribution = equal, Inequity Type = AI, Condition = Control, Decider Age Group = 6&7, Decider Gender = female. Griffith score (for first model only) and RMET score were included as continuous predictors. Table also shows goodness-of-fit statistics.

	Model
Intercept	-2.49 (0.51)***
Distribution	0.46 (0.86)
Inequity Type	-1.66 (0.91)†
Condition	0.44 (0.59)
Decider Age Group	0.16 (0.69)
RMET Score	0.06 (0.26)
Griffith Score	-0.24 (0.26)
Decider Gender	0.10 (0.23)
Distribution x Inequity Type	4.28 (1.33)**
Distribution x Condition	1.50 (1.08)
Distribution x Decider Age Group	1.09 (1.22)
Inequity Type x Condition	1.96 (1.05)†
Inequity Type x Decider Age Group	1.62 (1.14)
Condition x Decider Age Group	-0.46 (0.87)
Distribution x RMET Score	0.81 (0.44)†
Inequity Type x RMET Score	-0.09 (0.35)
Distribution x Griffith Score	0.54 (0.42)
Inequity Type x Griffith Score	0.20 (0.31)
Decider Age Group x RMET Score	0.05 (0.25)
Decider Age Group x Griffith Score	0.33 (0.22)
Distribution x Inequity Type x Condition	-4.05 (1.64)*
Distribution x Inequity Type x Decider Age Group	-2.87 (1.79)
Distribution x Condition x Decider Age Group	-2.11 (1.55)
Inequity Type x Condition x Decider Age Group	-1.58 (1.36)
Distribution x Inequity Type x RMET Score	-0.33 (0.61)
Distribution x Inequity Type x Griffith Score	-0.91 (0.56)
Distribution x Inequity Type x Condition x Decider Age Group	4.14 (2.23)†
AIC	1405.49
BIC	1564.40
Log Likelihood	-672.74

Num. obs. (trials)	1476
Num. groups: ID (subjects)	123
Variance: ID (Intercept)	0.93
Variance: ID (Distribution)	5.27
Covariance: ID (Intercept) (Distribution)	-1.60

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$

4.8.3 Supplemental results: With age as a continuous variable

In the main text, age was entered into the models as a categorical variable. When we entered age as a continuous variable, we found the same pattern of results:

There was no significant four-way interaction between Distribution, Inequity Type, Condition, and Decider Age Group, no significant three-way interaction between Distribution, Inequity Type, and Condition, and no significant two-way interaction between Distribution and Condition (four-way: LRT, $\chi^2_1 = 1.08$, $p = 0.30$; three-way: LRT, $\chi^2_1 = 2.65$, $p = 0.10$; two-way: LRT, $\chi^2_1 = 0.59$, $p = 0.44$). These results echo what we found in the main text: the control and theory of mind conditions did not differentially affect the rate of rejection for equal and unequal trials.

Similarly, there was no three-way interaction between Distribution, Inequity Type, and RMET score ($\chi^2_1 = 0.144$, $p = 0.70$), but there was a significant two-way interaction between Distribution and RMET score (LRT, $\chi^2_1 = 3.90$, p = 0.048). The pattern of results found here is the same as the one described in the main text (i.e., children higher in ToM are more likely to reject unequal distributions).

Here, too, we find a marginal three-way interaction between Distribution, Inequity Type, and Griffith score ($\chi^2_1 = 2.63$, $p = 0.11$) and no significant two-way interaction between Distribution Type and Griffith score ($\chi^2_1 = 0.003$, $p = 0.96$). These results, again, provide some weak evidence that the interaction between inequity type and Griffith score differed for advantageous inequity and disadvantageous inequity.

4.8.4 Supplemental results: Interactions with RMET score

Because GLMMs do not allow for missing data as predictor values, we were limited to data from 123 out of the 159 deciders (36 parents of deciders opted to not fill out or did not completely fill out the Griffith Empathy Measure; 3 of these deciders did not complete the RMET). By removing interactions with the Griffith score from the model, we were able to test the interaction between Distribution and RMET score on the entire sample of participants with valid RMET scores (N = 156).

The new model included the following predictor variables: Distribution (equal vs. unequal), Inequity Type (advantageous vs. disadvantageous), Condition (perspective-taking vs. control), Decider Age Group (6- to 7-year-olds vs. 8- to 9-year-olds), Decider Gender (female vs. male), and RMET score. A three-way interaction between Distribution, Inequity Type, and RMET score and two-way interactions between (1) Distribution and Inequity Type and (2) Distribution and RMET score were also included. We also included the two-way interaction between Decider Age Group and RMET score to control for possible effects of age. Participant ID was included as a random effect, and we fit a random slope and intercept for each Participant, allowing the intercept and effect of Distribution to vary across individuals.

While there was no three-way interaction between Distribution, Inequity Type, and RMET score, there was a significant two-way interaction between Distribution and RMET score (LRT, $X^2(1) = 3.91$, $p = 0.048$; Figure S1; Table S3). The pattern of results found here is similar to the one described in the main text (i.e., children higher in ToM are more likely to reject unequal distributions), albeit weaker.

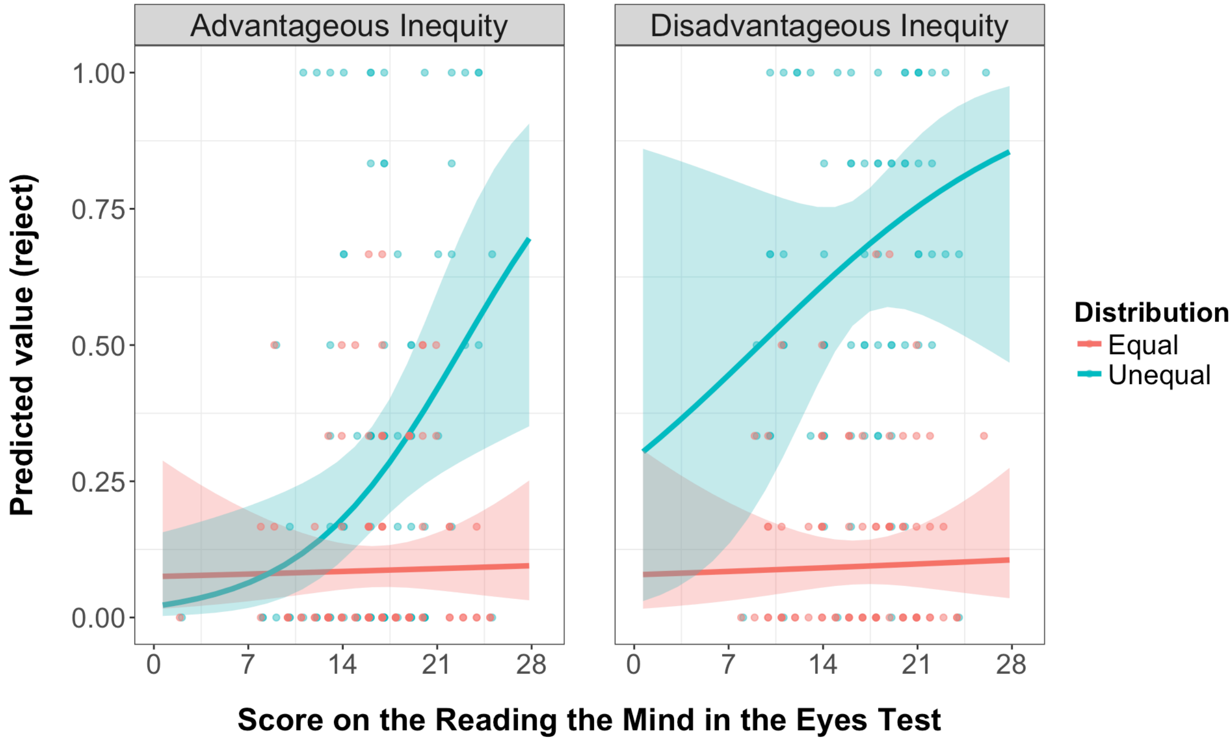


Figure S1. Rejections to equal and unequal distributions by score on the Reading the Mind in the Eyes Test. Lines and ribbons denote predicted values and confidence intervals extracted from the new model excluding predictors involving the Griffith Empathy Measure. Dots denote raw data.

Table S3. Model outputs from a reduced model (N=156) that includes all the interactions of interest related to RMET score in bold (Distribution x RMET score, Distribution x Inequity Type x RMET score). Estimates and standard errors (s.e.) of fixed effects in Generalized Linear Mixed Models predicting children's rejection behaviors. Baselines were set as follows: Distribution = equal, Inequity Type = AI, Decider Age Group = 6&7. RMET score was included as a continuous predictor. Table also shows goodness-of-fit statistics.

	Model
Intercept	-2.43 (0.26)***
Distribution	1.28 (0.38)***
Inequity Type	0.09 (0.29)
RMET Score	0.02 (0.23)
Decider Age Group	0.16 (0.22)
Distribution x Inequity Type	1.68 (0.52)**
Distribution x RMET Score	**0.68 (0.37)**†
Inequity Type x RMET Score	-0.04 (0.29)
Decider Age Group x RMET Score	0.10 (0.22)
Distribution x Inequity Type x RMET Score	**-0.30 (0.53)**
AIC	1766.91
BIC	1838.87
Log Likelihood	-870.46
Num. obs. (trials)	1872
Num. groups: ID (subjects)	156
Variance: ID (Intercept)	1.30
Variance: ID (Distribution)	6.80
Covariance: ID (Intercept) (Distribution)	-2.27

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.1$

5.0 GENERAL DISCUSSION

This dissertation presents three papers that investigate whether people display a sensitivity to context when engaging in theory of mind, with a large focus on cooperative and competitive contexts. Paper 1 presents evidence of better theory of mind for competition than cooperation among preschool children: preschool children are better at understanding false beliefs in competitive than cooperative contexts and better at applying their understanding in competitive interactions than cooperative interactions. Paper 2 presents no evidence of greater engagement of theory of mind for competition vs. cooperation or vice versa but it does provide evidence that brain regions implicated in theory of mind encode information about cooperation or competition. Moreover, it presents evidence showing links between brain and behavior that is context-dependent: activity in theory of mind regions during exposure to people's competitive, cooperative, and neutral actions differentially tracked with performance in the task and post-scan impressions of people in the vignettes. Paper 3 provides evidence showing a link between individual differences in theory of mind and responses to inequity but does not provide evidence of any differences in theory of mind across forms of inequity. Altogether, these results show an influence of theory of mind on social evaluations and social behaviors and support the idea that sensitivity to context may emerge early in life but becomes more difficult to detect over time.

There are three themes that connect these papers together. First is the use of multimodal approaches to ask similar questions about theory of mind across different social contexts. Second is the attempt made to understand the relationship between theory of mind and interactions across different social contexts—whether there is a link (Papers 1 and 3) and to what extent it is about the context (Paper 2). Third is the attention paid to individual differences in theory of mind.

The second theme is especially important given the lack of research on this very topic—a large focus of theory of mind research examines how theory of mind relates to moral and social judgment, but judgment and behavior does not necessarily correspond well with each other. In fact, even behaviors in hypothesized behaviors has been shown to differ from real behaviors (FeldmanHall, Dalgleish, et al., 2012; FeldmanHall, Mobbs, et al., 2012). While developmental research has led the way in examining relationships between theory of mind and actual social behaviors (e.g., prosocial acts; lying), this type of question has been examined to a lesser extent in adults. Why exactly this is the case is unclear. One potential concern is that it is difficult to tease apart influences of theory of mind from other influences (e.g., socialization, language, executive functioning) when examining people's behaviors, but this type of concern is not specific to behavior, also applies to judgments, and is certainly not limited to adults.

The third theme on individual differences is one that needs to be emphasized more in research. Assumptions that people have similar psychologies across the world have been debunked in domains such as fairness, cooperation, moral reasoning, and even in domains like visual perception (Henrich, Heine, & Norenzayan, 2010). How individual differences in theory of mind capacities relate to judgment and behavior would be a fruitful avenue of research. In Paper 3, we show that children who score higher on a theory of mind task tend to have greater rates of rejection to unequal distributions of resources. This is typically how individual differences are measured, but one could imagine that instead of examining differences in behaviors, one could also assess individual differences in the ability to produce the same behavior. For example, many different people can achieve the same score (e.g., 90) on an exam: the person who studied really hard, the person who did not study at all, and the overachiever who happened to be sick on the day of the exam and *only* scored a 90. In the context of cooperative behaviors, one could imagine people with different levels of theory of mind understanding responding the same way to one scenario (e.g., giving the same amount in a Dictator Game). One could imagine that direct manipulations of theory of mind could affect people's future giving behaviors differently depending on their current level of theory of mind understanding. A similar method of assessing and testing individual differences could potentially lead to personalized approaches to improving theory of mind capacities in individuals with impaired theory of mind, similar to personalized treatment in medicine.

Ongoing and future work

One aim of ongoing work is to examine the types of information encoded by regions implicated in theory of mind. In a prior study, participants played a game in which they had to cooperate or compete with the same person (in actuality, they played with a computer) (Tsoi, Dungan, Waytz, & Young, 2016). This study revealed that information separating cooperation from competition was encoded in theory of mind regions when people's responses in the game affected monetary outcomes. An extension of this work used eye-tracking to examine the types of information people are overtly attending to when playing this type of game. Preliminary results suggest that people spend similar amounts of time looking at their partner's screen (e.g., the partner's goal, the partner's response, the partner's monetary outcome) across cooperation and competition, which does not provide support for the idea that people are more motivated to attend to information pertaining to the partner during competition than during cooperation. A second extension is an ongoing study examining whether adults with ASD show similar encoding of context in theory of mind regions. We expect to find no encoding of context by theory of mind regions in individuals with ASD. Paper 2 was another extension of the fMRI study and examined whether information about context is also encoded when people are merely reading about people's behaviors in different contexts. Combined with the fMRI study above, it appears that information separating cooperation from competition may only be encoded in theory of mind regions when the information is highly relevant to the participant or the task itself.

A second aim of ongoing work is to examine how theory of mind changes across contexts *and* across different people. This work tries to tease apart effects of group membership and effects of context in theory of mind. In one initial study, we focused on altruistic behaviors as assessed by allocations in the Dictator Game. There was a perspective-taking manipulation, in which participants, prior to the game, had to write what they imagined the recipient would think and feel about the participant's decision (Perspective-Taking) or they had to write about something unrelated (Control). Before the task, participants were assigned to groups using a minimal group paradigm, and participants were told that they were either playing the Dictator Game with an ingroup member or an outgroup member. After the task, participants filled out a questionnaire assessing ingroup identification. In the Control condition, participants allocated similar amounts to ingroup members and outgroup members, and this did not differ across different levels of ingroup identification, whereas in the Perspective-Taking condition, participants with lower levels of ingroup identification gave more to outgroup members than to ingroup members. This result suggests that perspective-taking manipulations may work to influence people's altruistic behaviors, but it depends on individual differences in ingroup identification.

Altogether, the studies in this dissertation as well as ongoing work, serve to provide a starting point for many different investigations that combine theory of mind with social behaviors and intergroup and interpersonal processes, both in adults and in children.

6.0 REFERENCES

Agresti, A., & Coull, B. A. (1998). Approximate is better than "exact" for interval estimation of binomial proportions. *The American Statistician*, *52*(2), 119–126.

Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Do visual perspective tasks need theory of mind? *NeuroImage*, *30*(3), 1059–1068. https://doi.org/10.1016/j.neuroimage.2005.10.026

Aloise-Young, P. A. (1993). The development of self-presentation: Self-promotion in 6- to 10-year-old children. *Social Cognition*, *11*(2), 201–222. https://doi.org/10.1521/soco.1993.11.2.201

Andreoni, J., & Bernheim, B. D. (2009). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, *77*(5), 1607–1636. https://doi.org/10.3982/ECTA7384

Apfelbaum, E. P., Pauker, K., Ambady, N., Sommers, S. R., & Norton, M. I. (2008). Learning (not) to talk about race: When older children underperform in social categorization. *Developmental Psychology*, *44*(5), 1513–1518. https://doi.org/10.1037/a0012835

Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: A 1-year longitudinal perspective. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00655

Banerjee, R. (2002). Audience effects on self-presentation in childhood. *Social Development*, *11*(4), 487–507. https://doi.org/10.1111/1467-9507.00212

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*(2), 241–251.

Baron-Cohen, Simon. (2000). Theory of mind and autism: A fifteen year review. *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience (2nd Ed.)*, 3–20.

Baron-Cohen, Simon, Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, *21*(1), 37–46.

Baron-Cohen, Simon, Wheelwright, S., Spong, A., Scahill, V., & Lawson, J. (2001). Are intuitive physics and intuitive psychology independent? A test with children with Asperger Syndrome. *Journal of Developmental and Learning Disorders*, *5*(1), 47–78.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015a). Fitting linear mixed-effects models using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015b). Fitting linear mixed-effects models using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Batson, C. D., Thompson, E. R., & Chen, H. (2002). Moral hypocrisy: Addressing some alternatives. *Journal of Personality and Social Psychology*, *83*(2), 330–339. https://doi.org/10.1037/0022-3514.83.2.330

Baumard, N., Mascaro, O., & Chevallier, C. (2012). Preschoolers are able to take merit into account when distributing goods. *Developmental Psychology*, *48*(2), 492–498. https://doi.org/10.1037/a0026598

Begeer, S., Rieffe, C., Terwogt, M. M., & Stockmann, L. (2003). Theory of mind–based action in children from the autism spectrum. *Journal of Autism and Developmental Disorders*, *33*(5), 479–487.

Blake, P. R., McAuliffe, K., Corbit, J., Callaghan, T. C., Barry, O., Bowie, A., … Warneken, F. (2015). The ontogeny of fairness in seven societies. *Nature*, *528*(7581), 258–261. https://doi.org/10.1038/nature15703

Blake, Peter R., & McAuliffe, K. (2011). "I had so much it didn't seem fair": Eight-year-olds reject two forms of inequity. *Cognition*, *120*(2), 215–224. https://doi.org/10.1016/j.cognition.2011.04.006

Blake, Peter R., McAuliffe, K., & Warneken, F. (2014). The developmental origins of fairness: the knowledge–behavior gap. *Trends in Cognitive Sciences*, *18*(11), 559–561. https://doi.org/10.1016/j.tics.2014.08.003

Blake, Peter R., Piovesan, M., Montinari, N., Warneken, F., & Gino, F. (2015). Prosocial norms in the classroom: The role of self-regulation in following norms of giving. *Journal of Economic Behavior & Organization*, *115*, 18–29. https://doi.org/10.1016/j.jebo.2014.10.004

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, *77*(1), B25-31.

Brosnan, S. F., & Bshary, R. (2016). On potential links between inequity aversion and the structure of interactions for the evolution of cooperation. *Behaviour*, *153*(9–11), 1267–1292. https://doi.org/10.1163/1568539X-00003355

Bruneau, E. G., & Saxe, R. (2012). The power of being heard: The benefits of 'perspective-giving' in the context of intergroup conflict. *Journal of Experimental Social Psychology*, *48*(4), 855–866. https://doi.org/10.1016/j.jesp.2012.02.017

Buhrmester, D., Goldfarb, J., & Cantrell, D. (1992). Self-presentation when sharing with friends and nonfriends. *The Journal of Early Adolescence*, *12*(1), 61–79.

Bull, R., Phillips, L. H., & Conway, C. A. (2008). The role of control functions in mentalizing: Dual-task studies of Theory of Mind and executive function. *Cognition*, *107*(2), 663–672. https://doi.org/10.1016/j.cognition.2007.07.015

Burkart, J. M., Hrdy, S. B., & Van Schaik, C. P. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, *18*(5), 175–186. https://doi.org/10.1002/evan.20222

Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language*, *28*(5), 606–637. https://doi.org/10.1111/mila.12036

Byrne, R. W., & Corp, N. (2004). Neocortex size predicts deception rate in primates. *Proceedings of the Royal Society B: Biological Sciences*, *271*(1549), 1693–1699. https://doi.org/10.1098/rspb.2004.2780

Byrne, Richard W., & Whiten, A. (Eds.). (1988). *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford : New York: Clarendon Press ; Oxford University Press.

Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., & Eickhoff, S. B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure and Function*, *217*(4), 783–796. https://doi.org/10.1007/s00429-012-0380-y

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*(4), 1032–1053.

Carlson, Stephanie M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development*, *11*(2), 73–92. https://doi.org/10.1002/icd.298

Carlson, Stephanie M, Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, *87*(4), 299–319. https://doi.org/10.1016/j.jecp.2004.01.002

Carruthers, P. (2016). Two Systems for Mindreading? *Review of Philosophy and Psychology*, *7*(1), 141–162. https://doi.org/10.1007/s13164-015-0259-y

Chandler, M., Fritz, A. S., & Hala, S. (1989). Small-Scale Deceit: Deception as a Marker of Two-, Three-, and Four-Year-Olds' Early Theories of Mind. *Child Development*, *60*(6), 1263. https://doi.org/10.2307/1130919

Christensen, R. H. B. (2015). *ordinal—Regression Models for Ordinal Data*.

Cikara, M., Bruneau, E. G., & Saxe, R. R. (2011). Us and Them: Intergroup Failures of Empathy. *Current Directions in Psychological Science*, *20*(3), 149–153. https://doi.org/10.1177/0963721411408713

Corbit, J., McAuliffe, K., Callaghan, T. C., Blake, P. R., & Warneken, F. (2017). Children's collaboration induces fairness rather than generosity. *Cognition*, *168*, 344–356. https://doi.org/10.1016/j.cognition.2017.07.006

Cortes, B. P., Demoulin, S., Rodriguez, R. T., Rodriguez, A. P., & Leyens, J.-P. (2005). Infrahumanization or Familiarity? Attribution of Uniquely Human Emotions to the Self, the Ingroup, and the Outgroup. *Personality and Social Psychology Bulletin*, *31*(2), 243–253. https://doi.org/10.1177/0146167204271421

Dadds, M. R., Hunter, K., Hawes, D. J., Frost, A. D. J., Vassallo, S., Bunn, P., … Masry, Y. E. (2008). A measure of cognitive and affective empathy in children using parent ratings. *Child Psychiatry and Human Development*, *39*(2), 111–122. https://doi.org/10.1007/s10578-007-0075-4

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, *33*(1), 67–80. https://doi.org/10.1007/s00199-006-0153-z

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. *Nature*, *446*(7137), 794–796. https://doi.org/10.1038/nature05651

Decety, J., & Cacioppo, S. (2012). The speed of morality: A high-density electrical neuroimaging study. *Journal of Neurophysiology*, *108*(11), 3068–3072. https://doi.org/10.1152/jn.00473.2012

Decety, Jean, Jackson, P. L., Sommerville, J. A., Chaminade, T., & Meltzoff, A. N. (2004). The neural bases of cooperation and competition: An fMRI investigation. *NeuroImage*, *23*(2), 744–751. https://doi.org/10.1016/j.neuroimage.2004.05.025

Devine, R. T., & Hughes, C. (2016). Measuring theory of mind across middle childhood: Reliability and validity of the Silent Films and Strange Stories tasks. *Journal of Experimental Child Psychology*, *149*, 23–40. https://doi.org/10.1016/j.jecp.2015.07.011

Devine, R. T., White, N., Ensor, R., & Hughes, C. (2016). Theory of mind in middle childhood: Longitudinal associations with executive function and social competence. *Developmental Psychology*, *52*(5), 758–771. https://doi.org/10.1037/dev0000105

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *NeuroImage*, *55*(2), 705–712. https://doi.org/10.1016/j.neuroimage.2010.12.040

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, *25*(4), 1325–1335. https://doi.org/10.1016/j.neuroimage.2004.12.034

Engelmann, J. M., Herrmann, E., & Tomasello, M. (2012). Five-year olds, but not chimpanzees, attempt to manage their reputations. *PLoS ONE*, *7*(10), e48433. https://doi.org/10.1371/journal.pone.0048433

Engelmann, J. M., & Rapp, D. J. (2018). The influence of reputational concerns on children's prosociality. *Current Opinion in Psychology*, *20*, 92–95. https://doi.org/10.1016/j.copsyc.2017.08.024

Epley, N., Akalis, S., Waytz, A., & Cacioppo, J. T. (2008). Creating Social Connection Through Inferential Reproduction: Loneliness and Perceived Agency in Gadgets, Gods, and Greyhounds. *Psychological Science*, *19*(2), 114–120. https://doi.org/10.1111/j.1467-9280.2008.02056.x

Epley, N., & Waytz, A. (2010). Mind Perception. In *Handbook of Social Psychology*. John Wiley & Sons, Inc. Retrieved from http://dx.doi.org/10.1002/9780470561119.socpsy001014

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868.

FeldmanHall, O., Dalgleish, T., Thompson, R., Evans, D., Schweizer, S., & Mobbs, D. (2012). Differential neural circuitry and self-interest in real *vs* hypothetical moral decisions. *Social Cognitive and Affective Neuroscience*, *7*(7), 743–751. https://doi.org/10.1093/scan/nss069

FeldmanHall, O., Mobbs, D., Evans, D., Hiscox, L., Navrady, L., & Dalgleish, T. (2012). What we say and what we do: The relationship between real and hypothetical moral choices. *Cognition*, *123*(3), 434–441. https://doi.org/10.1016/j.cognition.2012.02.001

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly*, *50*(3), 274–290.

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of “theory of mind” in story comprehension. *Cognition*, *57*(2), 109–128.

Flombaum, J. I., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology*, *15*(5), 447–452. https://doi.org/10.1016/j.cub.2004.12.076

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*(4), 708–724.

Galinsky, Adam D., Maddux, W. W., Gilin, D., & White, J. B. (2008). Why it pays to get inside the head of your opponent: The differential effects of perspective taking and empathy in negotiations. *Psychological Science*, *19*(4), 378–384. https://doi.org/10.1111/j.1467-9280.2008.02096.x

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of “theory of mind” in verbal and nonverbal tasks. *Neuropsychologia*, *38*(1), 11–21.

Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants’ reactions to equal and unequal distributions of resources. *Developmental Science*, *14*(5), 1012–1020. https://doi.org/10.1111/j.1467-7687.2011.01048.x

Gobbini, M. I., Koralek, A. C., Bryan, R. E., Montgomery, K. J., & Haxby, J. V. (2007). Two takes on the social brain: A comparison of theory of mind tasks. *Journal of*

Cognitive Neuroscience, *19*(11), 1803–1814. https://doi.org/10.1162/jocn.2007.19.11.1803

Gonzalez, Antonya M., Steele, J. R., & Baron, A. S. (2017). Reducing children's implicit racial bias through exposure to positive out-group exemplars. *Child Development*, *88*(1), 123–130. https://doi.org/10.1111/cdev.12582

Gonzalez, Antonya Marie, Dunlop, W. L., & Baron, A. S. (2016). Malleability of implicit associations across development. *Developmental Science*. https://doi.org/10.1111/desc.12481

Greene, Joshua D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105–2108.

Greene, Joshua David. (2013). *Moral tribes: Emotion, reason, and the gap between us and them*. New York: The Penguin Press.

Hadwin, J., & Perner, J. (1991). Pleased and surprised: Children's cognitive theory of emotion. *British Journal of Developmental Psychology*, *9*(2), 215–234. https://doi.org/10.1111/j.2044-835X.1991.tb00872.x

Hala, S., Chandler, M., & Fritz, A. S. (1991). Fledgling theories of mind: Deception as a marker of three-year-olds' understanding of false belief. *Child Development*, *62*(1), 83–97. https://doi.org/10.1111/j.1467-8624.1991.tb01516.x

Hamann, K., Bender, J., & Tomasello, M. (2014). Meritocratic sharing is based on collaboration in 3-year-olds. *Developmental Psychology*, *50*(1), 121–128. https://doi.org/10.1037/a0032965

Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature*, *476*(7360), 328–331. https://doi.org/10.1038/nature10278

Happé, F. G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, *66*(3), 843–855.

Hare, B. (2001). Can competitive paradigms increase the validity of experiments on primate social cognition? *Animal Cognition*, *4*(3–4), 269–280. https://doi.org/10.1007/s100710100084

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*(4), 771–785. https://doi.org/10.1006/anbe.1999.1377

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, *61*(1), 139–151. https://doi.org/10.1006/anbe.2000.1518

Hare, B., Call, J., & Tomasello, M. (2006). Chimpanzees deceive a human competitor by hiding. *Cognition*, *101*(3), 495–514. https://doi.org/10.1016/j.cognition.2005.01.011

Hare, B., & Tomasello, M. (2004). Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. *Animal Behaviour*, *68*(3), 571–581. https://doi.org/10.1016/j.anbehav.2003.11.011

Harris, L. T., & Fiske, S. T. (2007). Social groups that elicit disgust are differentially processed in mPFC. *Social Cognitive and Affective Neuroscience*, *2*(1), 45–51. https://doi.org/10.1093/scan/nsl037

Harris, L. T., & Fiske, S. T. (2009). Social neuroscience evidence for dehumanised perception. *European Review of Social Psychology*, *20*(1), 192–231. https://doi.org/10.1080/10463280902954988

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, *10*(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4

Haslam, N., & Loughnan, S. (2014). Dehumanization and infrahumanization. *Annual Review of Psychology*, *65*(1), 399–423. https://doi.org/10.1146/annurev-psych-010213-115045

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, *57*(2), 243. https://doi.org/10.2307/1416950

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Heyes, C. M. (1998). Theory of mind in nonhuman primates. *The Behavioral and Brain Sciences*, *21*(1), 101–114; discussion 115-148.

Hrdy, S. B. (2009). *Mothers and others: The evolutionary origins of mutual understanding*. Cambridge, Mass: Belknap Press of Harvard University Press.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology*, *16*(2), 233–253. https://doi.org/10.1111/j.2044-835X.1998.tb00921.x

Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development*, *13*(4), 590–619.

Jackson, J. W. (1993). Realistic group conflict theory: A review and evaluation of the theoretical and empirical literature. *The Psychological Record*, *43*(3), 395–413.

Jenkins, A. C., & Mitchell, J. P. (2010). Mentalizing under uncertainty: Dissociated neural responses to ambiguous and unambiguous mental state inferences. *Cerebral Cortex*, *20*(2), 404–410. https://doi.org/10.1093/cercor/bhp109

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Kahneman, D. (2012). *Thinking, fast and slow*. London: Penguin Books.

Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition*, *109*(2), 224–234. https://doi.org/10.1016/j.cognition.2008.08.010

Kanngiesser, P., & Warneken, F. (2012). Young children consider merit when sharing resources with others. *PLoS ONE*, *7*(8), e43979. https://doi.org/10.1371/journal.pone.0043979

Karg, K., Schmelz, M., Call, J., & Tomasello, M. (2015). The goggles experiment: can chimpanzees use self-experience to infer what a competitor can see? *Animal Behaviour*, *105*, 211–221. https://doi.org/10.1016/j.anbehav.2015.04.028

Kelman, H. G. (1973). Violence without moral restraint: Reflections on the dehumanization of victims and victimizers. *Journal of Social Issues*, *29*(4), 25–61. https://doi.org/10.1111/j.1540-4560.1973.tb00102.x

Koster-Hale, J., Saxe, R., Dungan, J., & Young, L. L. (2013). Decoding moral judgments from neural representations of intentions. *Proceedings of the National Academy of Sciences*, *110*(14), 5648–5653. https://doi.org/10.1073/pnas.1207992110

Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology*, *90*(4), 543–555. https://doi.org/10.1037/0022-3514.90.4.543

Lee, K. (2013). Little Liars: Development of Verbal Deception in Children. *Child Development Perspectives*, *7*(2), 91–96. https://doi.org/10.1111/cdep.12023

Leimgruber, K. L., Shaw, A., Santos, L. R., & Olson, K. R. (2012). Young children are more generous when others are aware of their actions. *PLoS ONE*, *7*(10), e48292. https://doi.org/10.1371/journal.pone.0048292

Leyens, J.-P., Paladino, P. M., Rodriguez-Torres, R., Vaes, J., Demoulin, S., Rodriguez-Perez, A., & Gaunt, R. (2000). The emotional side of prejudice: The attribution of secondary emotions to ingroups and outgroups. *Personality and Social Psychology Review*, *4*(2), 186–197. https://doi.org/10.1207/S15327957PSPR0402_06

Lissek, S., Peters, S., Fuchs, N., Witthaus, H., Nicolas, V., Tegenthoff, M., … Brüne, M. (2008). Cooperation and deception recruit different subsets of the theory-of-mind network. *PLoS ONE*, *3*(4), e2023. https://doi.org/10.1371/journal.pone.0002023

Liu, B., Huang, Z., Xu, G., Jin, Y., Chen, Y., Li, X., … Jing, J. (2016). Altruistic sharing behavior in children: Role of theory of mind and inhibitory control. *Journal of Experimental Child Psychology*, *141*, 222–228. https://doi.org/10.1016/j.jecp.2015.09.010

Loewenstein, G. F., Thompson, L., & Bazerman, M. H. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social Psychology*, *57*(3), 426.

Lurz, R. W. (2011). *Mindreading animals: The debate over what animals know about other minds*. Cambridge, Mass: MIT Press.

Lyons, D. E., & Santos, L. R. (2006). Ecology, domain specificity, and the origins of theory of mind: Is competition the catalyst? *Philosophy Compass*, *1*(5), 481–492. https://doi.org/10.1111/j.1747-9991.2006.00032.x

MacLaren, R., & Olson, D. (1993). Trick or treat: Children's understanding of surprise. *Cognitive Development*, *8*(1), 27–46. https://doi.org/10.1016/0885-2014(93)90003-N

Malle, B. F., & Holbrook, J. (2012). Is there a hierarchy of social inferences? The likelihood and speed of inferring intentionality, mind, and personality. *Journal of Personality and Social Psychology*, *102*(4), 661–684. https://doi.org/10.1037/a0026790

McAuliffe, K., Blake, P. R., Kim, G., Wrangham, R. W., & Warneken, F. (2013). Social influences on inequity aversion in children. *PLoS ONE*, *8*(12), e80966. https://doi.org/10.1371/journal.pone.0080966

McAuliffe, K., Blake, P. R., Steinbeis, N., & Warneken, F. (2017). The developmental foundations of human fairness. *Nature Human Behaviour*, *1*, 0042.

McAuliffe, K., & Dunham, Y. (2016). Group bias in cooperative norm enforcement. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150073. https://doi.org/10.1098/rstb.2015.0073

McPherson Frantz, C., & Janoff-Bulman, R. (2000). Considering Both Sides: The Limits of Perspective Taking. *Basic and Applied Social Psychology*, *22*(1), 31–42. https://doi.org/10.1207/S15324834BASP2201_4

Melis, A. P., Call, J., & Tomasello, M. (2006). Chimpanzees (Pan troglodytes) conceal visual and auditory information from others. *Journal of Comparative Psychology*, *120*(2), 154–162. https://doi.org/10.1037/0735-7036.120.2.154

Morewedge, C. K. (2009). Negativity bias in attribution of external agency. *Journal of Experimental Psychology: General*, *138*(4), 535–545. https://doi.org/10.1037/a0016796

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science (New York, N.Y.)*, *308*(5719), 255–258. https://doi.org/10.1126/science.1107621

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. *Frontiers in Neuroinformatics*, *10*. https://doi.org/10.3389/fninf.2016.00027

Opotow, S. (1990). Moral exclusion and injustice: An introduction. *Journal of Social Issues*, *46*(1), 1–20. https://doi.org/10.1111/j.1540-4560.1990.tb00268.x

Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology*, *1*(1), 33–60.

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental Psychology*, *43*(4), 974–990. https://doi.org/10.1037/0012-1649.43.4.974

Perner, J, & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences*, *3*(9), 337–344.

Perner, Josef, & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences*, *16*(10), 519–525. https://doi.org/10.1016/j.tics.2012.08.004

Peterson, C. C., Slaughter, V., Peterson, J., & Premack, D. (2013). Children with autism can track others' beliefs in a competitive game. *Developmental Science*, *16*(3), 443–450. https://doi.org/10.1111/desc.12040

Phillips, W., Barnes, J. L., Mahajan, N., Yamaguchi, M., & Santos, L. R. (2009). 'Unwilling' versus 'unable': capuchin monkeys' (*Cebus apella*) understanding of human intentional action: Unwilling vs. unable in capuchin monkeys. *Developmental Science*, *12*(6), 938–945. https://doi.org/10.1111/j.1467-7687.2009.00840.x

Povinelli, D. J., Rulf, A. B., & Bierschwale, D. T. (1994). Absence of knowledge attribution and self-recognition in young chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *108*(1), 74–80. https://doi.org/10.1037/0735-7036.108.1.74

Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: suspiciously human? *Trends in Cognitive Sciences*, *7*(4), 157–160. https://doi.org/10.1016/S1364-6613(03)00053-6

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(04), 515. https://doi.org/10.1017/S0140525X00076512

Raihani, N. J., & McAuliffe, K. (2012). Human punishment is motivated by inequity aversion, not a desire for reciprocity. *Biology Letters*, *8*(5), 802–804. https://doi.org/10.1098/rsbl.2012.0470

Reis, H. T., & Gruzen, J. (1976). On mediating equity, equality, and self-interest: The role of self-presentation in social exchange. *Journal of Experimental Social Psychology*, *12*(5), 487–503. https://doi.org/10.1016/0022-1031(76)90079-2

Rizzo, M. T., & Killen, M. (2016). Children's understanding of equity in the context of inequality. *British Journal of Developmental Psychology*, *34*(4), 569–581. https://doi.org/10.1111/bjdp.12150

Rochat, P., Dias, M. D. G., Guo Liping, Broesch, T., Passos-Ferreira, C., Winning, A., & Berg, B. (2009). Fairness in distributive justice by 3- and 5-year-olds across seven cultures. *Journal of Cross-Cultural Psychology*, *40*(3), 416–442. https://doi.org/10.1177/0022022109332844

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and U.S. preschoolers. *Psychological Science*, *17*(1), 74–81. https://doi.org/10.1111/j.1467-9280.2005.01667.x

Santiesteban, I., Banissy, M. J., Catmur, C., & Bird, G. (2015). Functional lateralization of temporoparietal junction - imitation inhibition, visual perspective-taking and theory of mind. *European Journal of Neuroscience*, *42*(8), 2527–2533. https://doi.org/10.1111/ejn.13036

Santos, L. R., Nissen, A. G., & Ferrugia, J. A. (2006). Rhesus monkeys, Macaca mulatta, know what others can and cannot hear. *Animal Behaviour*, *71*(5), 1175–1181. https://doi.org/10.1016/j.anbehav.2005.10.007

Saxe, R, & Kanwisher, N. (2003). People thinking about thinking peopleThe role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*(4), 1835–1842. https://doi.org/10.1016/S1053-8119(03)00230-1

Saxe, R, & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, *43*(10), 1391–1399. https://doi.org/10.1016/j.neuropsychologia.2005.02.013

Saxe, Rebecca, Carey, S., & Kanwisher, N. (2004). Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology*, *55*, 87–124. https://doi.org/10.1146/annurev.psych.55.090902.142044

Saxe, Rebecca, & Powell, L. J. (2006). It's the thought that counts: Specific brain regions for one component of theory of mind. *Psychological Science*, *17*(8), 692–699. https://doi.org/10.1111/j.1467-9280.2006.01768.x

Schmelz, M., Call, J., & Tomasello, M. (2011). Chimpanzees know that others make inferences. *Proceedings of the National Academy of Sciences*, *108*(7), 3077–3079. https://doi.org/10.1073/pnas.1000469108

Schmelz, Martin, & Call, J. (2016). The psychology of primate cooperation and competition: a call for realigning research agendas. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150067. https://doi.org/10.1098/rstb.2015.0067

Schmidt, M. F. H., & Sommerville, J. A. (2011). Fairness expectations and altruistic sharing in 15-month-old human infants. *PLoS ONE*, *6*(10), e23223. https://doi.org/10.1371/journal.pone.0023223

Schug, J., Takagishi, H., Benech, C., & Okada, H. (2016). The development of theory of mind and positive and negative reciprocity in preschool children. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00888

Schurz, M., Aichhorn, M., Martin, A., & Perner, J. (2013). Common brain areas engaged in false belief reasoning and visual perspective taking: a meta-analysis of functional brain imaging studies. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00712

Scott, R. M., & Baillargeon, R. (2017). Early False-Belief Understanding. *Trends in Cognitive Sciences*, *21*(4), 237–249. https://doi.org/10.1016/j.tics.2017.01.012

Shaw, A., Montinari, N., Piovesan, M., Olson, K. R., Gino, F., & Norton, M. I. (2014). Children develop a veil of fairness. *Journal of Experimental Psychology: General*, *143*(1), 363–375. https://doi.org/10.1037/a0031247

Sheskin, M., Bloom, P., & Wynn, K. (2014). Anti-equality: Social comparison in young children. *Cognition*, *130*(2), 152–156. https://doi.org/10.1016/j.cognition.2013.10.008

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience & Biobehavioral Reviews*, *30*(6), 855–863. https://doi.org/10.1016/j.neubiorev.2006.06.011

Singer, T. (2009). Understanding others: Brain mechanisms of theory of mind and empathy. *Neuroeconomics: Decision Making and the Brain*, 251–268.

Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness? *Psychological Science*, *23*(2), 196–204. https://doi.org/10.1177/0956797611422072

Smith, C. E., Blake, P. R., & Harris, P. L. (2013). I should but I won't: Why young children endorse norms of fair sharing but do not follow them. *PloS One*, *8*(3), e59510. https://doi.org/10.1371/journal.pone.0059510

Struch, N., & Schwartz, S. H. (1989). Intergroup aggression: Its predictors and distinctness from in-group bias. *Journal of Personality and Social Psychology*, *56*(3), 364–373.

Takagishi, H., Koizumi, M., Fujii, T., Schug, J., Kameshima, S., & Yamagishi, T. (2014). The role of cognitive and emotional perspective taking in economic decision making in the Ultimatum Game. *PLoS ONE*, *9*(9), e108462. https://doi.org/10.1371/journal.pone.0108462

Thornton, A., & McAuliffe, K. (2015). Cognitive consequences of cooperative breeding? A critical appraisal: Cognitive consequences of cooperative breeding. *Journal of Zoology*, *295*(1), 12–22. https://doi.org/10.1111/jzo.12198

Todd, A. R., & Galinsky, A. D. (2014). Perspective-Taking as a Strategy for Improving Intergroup Relations: Evidence, Mechanisms, and Qualifications: Perspective-Taking and Intergroup Relations. *Social and Personality Psychology Compass*, *8*(7), 374–387. https://doi.org/10.1111/spc3.12116

Tsoi, L., Dungan, J. A., Chakroff, A., & Young, L. L. (2018). Neural substrates for moral judgments of psychological versus physical harm. *Social Cognitive and Affective Neuroscience*. https://doi.org/10.1093/scan/nsy029

Tsoi, L., Dungan, J., Waytz, A., & Young, L. (2016). Distinct neural patterns of social cognition for cooperation versus competition. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2016.04.069

Tsoi, L., Hamlin, J. K., Baron, A. S., Waytz, A., & Young, L. (under review). Developmental evidence for enhanced theory of mind for competitive versus cooperative contexts.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, *134*(3), 383–403. https://doi.org/10.1037/0033-2909.134.3.383

Valdesolo, P., & DeSteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion*, *11*(2), 262–266. https://doi.org/10.1037/a0021302

Wagner, D. D., Kelley, W. M., & Heatherton, T. F. (2011). Individual differences in the spontaneous recruitment of brain regions supporting mental state understanding when viewing natural social scenes. *Cerebral Cortex*, *21*(12), 2788–2796. https://doi.org/10.1093/cercor/bhr074

Waytz, A., Epley, N., & Cacioppo, J. T. (2010). Social Cognition Unbound: Insights Into Anthropomorphism and Dehumanization. *Current Directions in Psychological Science*, *19*(1), 58–62. https://doi.org/10.1177/0963721409359302

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, *14*(8), 383–388. https://doi.org/10.1016/j.tics.2010.05.006

Waytz, A., Morewedge, C. K., Epley, N., Monteleone, G., Gao, J.-H., & Cacioppo, J. T. (2010a). Making sense by making sentient: Effectance motivation increases anthropomorphism. *Journal of Personality and Social Psychology*, *99*(3), 410–435. https://doi.org/10.1037/a0020240

Waytz, A., Morewedge, C. K., Epley, N., Monteleone, G., Gao, J.-H., & Cacioppo, J. T. (2010b). Making sense by making sentient: Effectance motivation increases anthropomorphism. *Journal of Personality and Social Psychology*, *99*(3), 410–435. https://doi.org/10.1037/a0020240

Waytz, A., & Young, L. (2014). Two motivations for two dimensions of mind. *Journal of Experimental Social Psychology*, *55*, 278–283. https://doi.org/10.1016/j.jesp.2014.08.001

Wellman, H. M., & Banerjee, M. (1991). Mind and emotion: Children's understanding of the emotional consequences of beliefs and desires. *British Journal of Developmental Psychology*, *9*(2), 191–214. https://doi.org/10.1111/j.2044-835X.1991.tb00871.x

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655–684.

Wheatley, T., Milleville, S. C., & Martin, A. (2007). Understanding animate agents: Distinct roles for the social network and mirror system. *Psychological Science*, *18*(6), 469–474. https://doi.org/10.1111/j.1467-9280.2007.01923.x

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*(5), 297–333. https://doi.org/10.1037/h0040934

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and Cooperation. *Psychological Science*, *20*(1), 1–5. https://doi.org/10.1111/j.1467-9280.2008.02253.x

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128.

Woodruff, G., & Premack, D. (1979). Intentional communication in the chimpanzee: The development of deception. *Cognition*, *7*(4), 333–362. https://doi.org/10.1016/0010-0277(79)90021-0

Yirmiya, N., Erel, O., Shaked, M., & Solomonica-Levi, D. (1998). Meta-analyses comparing theory of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. *Psychological Bulletin*, *124*(3), 283–307.

Yoder, K. J., & Decety, J. (2014). The Good, the Bad, and the Just: Justice sensitivity predicts neural response during moral evaluation of actions performed by others. *Journal of Neuroscience*, *34*(12), 4161–4166. https://doi.org/10.1523/JNEUROSCI.4648-13.2014

Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the*

National Academy of Sciences, *104*(20), 8235–8240. https://doi.org/10.1073/pnas.0701408104

Young, Liane, & Saxe, R. (2009a). An fMRI investigation of spontaneous mental state inference for moral judgment. *Journal of Cognitive Neuroscience*, *21*(7), 1396–1405. https://doi.org/10.1162/jocn.2009.21137

Young, Liane, & Saxe, R. (2009b). Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia*, *47*(10), 2065–2072. https://doi.org/10.1016/j.neuropsychologia.2009.03.020

Zaki, J., & Ochsner, K. (2016). Empathy. In *Handbook of Emotion* (4th ed., pp. 871–884). New York: Guilford.