A cooperation advantage for theory of mind in children and adults

Lily Tsoi[1], Kiley Hamlin[2], Adam Waytz[3], Andrew Scott Baron[2], Liane Young[4]

[1] Department of Psychology, Princeton University

[2] Department of Psychology, University of British Columbia

[3] Management and Organizations Department, Northwestern University

[4] Department of Psychology, Boston College

Corresponding Author:
Lily Tsoi
Department of Psychology
Princeton University
Peretsman Scully Hall
Princeton, NJ 08540
Email: ltsoi@princeton.edu

Declarations of interest: None

Abstract

Three studies test whether people engage in mental state reasoning or Theory of Mind (ToM) differently across two fundamental social contexts: cooperation and competition. Study 1 examines how children with an emerging understanding of false beliefs deploy ToM across these contexts. We find that young preschool children are better able to plant false beliefs in others' minds in a cooperative versus competitive context; this difference does not emerge for other cognitive capacities tested (e.g., executive functioning, memory). Studies 2a and 2b reveal the same systematic difference in adults' ToM for cooperation and competition, even after accounting for relevant predictors (e.g., preference for a task condition; feelings about deception). Together, these findings provide initial evidence for enhanced ToM for cooperation versus competition in early development and also adulthood.

Keywords: social cognition; social development; moral behavior; cooperation; competition; false belief

Introduction

Cooperation and competition comprise two fundamental forms of social interaction. At first glance, successful cooperative and competitive interactions both appear to require the capacity to infer, attribute, and reason about the contents of agents' minds (e.g., thoughts, intentions, and beliefs), a capacity often referred to as theory of mind (ToM). Indeed, to successfully help another person, one must understand both that the person desires a particular outcome and that they are unable to achieve it. Similarly, to compete effectively against another person, one must understand what they are thinking in order to effectively oppose them.

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 (for a review, see Lyons & Santos, 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. There is some emerging evidence that chimpanzees can track others' mental states (e.g., what they can or cannot see) in cooperative contexts (Grueneisen, Duguid, Saur, & Tomasello, 2017), but for the most part the existing empirical record is in line with evolutionary theories proposing that ToM 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, perhaps in the service of understanding those agents' present and future behaviors (Hamlin & Baron, 2014; Knobe, n.d.; Leslie, Knobe, & Cohen, 2006; Morewedge, 2009; Vaish, Grossmann, & Woodward, 2008; Waytz et al., 2010).

By contrast, several lines of research suggest that, in humans, in particular, 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). The need to identify individuals in one’s environment who are most likely to provide optimal care may have facilitated the emergence of advanced ToM in humans. This idea is supported by the observation that, whereas other cooperative breeders in the primate lineage, most notably callitrichids, are largely deficient when compared to apes and various other non-cooperative breeding primates in general cognitive functioning, they nevertheless exhibit relatively advanced social-cognitive abilities (Burkart, Hrdy, & Van Schaik, 2009; but see Thornton & McAuliffe, 2015). Relatedly, experimental work in human children and adults reveals a greater tendency for people to consider the minds of ingroup versus outgroup members, perhaps because people are more likely to cooperate with ingroup members and compete with outgroup members (Kelman, 1973; Leyens et al., 2000; McLoughlin & Over, 2017; Opotow, 1990; Struch & Schwartz, 1989).

The current work builds on this rich history of theoretical and empirical work on the evolutionary origins or ToM as well as investigations of ToM in both human children and adults. We therefore tested whether children, at the point in development during which explicit ToM emerges, are better at deploying ToM for cooperation than competition or vice versa, and whether children and adults exhibited broadly similar patterns. Given that young children have arguably not been socialized to deploy their emerging ToM capacities differently across cooperative versus competitive contexts, and are just beginning to engage in mature, explicit ToM reasoning, our developmental approach may serve to shed light on whether an asymmetry in ToM reasoning is a fundamental feature of humans' reasoning about other minds.

We tested ToM during live social interactions and used 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). In Study 1, we examined whether 4-year-olds 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) than a cooperative goal (i.e., to be joint winners, together, of stickers) or vice versa. We focused on 4-year-olds, since at this age children typically have just begun to show explicit understanding of false beliefs, a key component of ToM involving the understanding that people can have beliefs that contradict reality (Wimmer & Perner, 1983; for a review, Wellman, Cross, & Watson, 2001). We did not test children younger than 4 years old because before this age children are typically unable to deceive others and require extensive training to systematically use deception to win a game (Ding, Heyman, Fu, Zhu, & Lee, 2018; Lee, 2013; our pilot tests). Moreover, because prior work has linked executive functioning with ToM (Carlson et al., 1998; Carlson & Moses, 2001; Carlson et al., 2002; Carlson et al., 2004; Gordon & Olson, 1998; Hughes, 1998), and response inhibition with strategic deception (Russell, Mauthner, Sharpe, & Tidswell, 1991), we also examined whether any difference in ToM across cooperation versus competition is specific to ToM or can instead be attributed to differences in memory or executive functioning across contexts. Participants therefore answered a memory question about the ToM task and also completed a cooperative or competitive version of a child-friendly response inhibition task (Day-Night task; Gerstadt, Hong, & Diamond, 1994).

In Studies 2a and 2b, we examined performance on the same ToM task in adults and probed, in a series of exploratory analyses, whether any difference in ToM across cooperation versus competition could be explained by any of the following: (1) people putting more effort into the game in one context versus the other, as indexed by response time, (2) people liking the game more in one context versus the other, (3) people preferring the person with whom they played the cooperative versus competitive game or vice versa, or (4) people feeling worse about planting a false belief in others (e.g., deceiving) in one context versus the other.

How people navigate these fundamental social contexts of cooperation and competition is a key question for scholars of moral psychology and intergroup cognition. The present research recognizes that identifying the moral status of agents, as friend or foe, for example, can powerfully impact social cognition, both in development and at the mature state. We therefore treat this core component of moral psychology—how people think about others across cooperative and competitive contexts—as a central hub connecting social cognition and developmental psychology to perspectives in evolutionary psychology and intergroup cognition. In particular, and informed by relevant work in evolutionary psychology (Lyons & Santos, 2006) and our own work in social psychology and social neuroscience (Tsoi & Young, 2018), we present three studies targeting how people deploy social cognitive capacities across cooperation and competition, using methods from social cognition and developmental psychology.

Study 1

We examined whether preschool children, in live interactions with an experimenter, are better able to plant false beliefs in the experimenter's mind to achieve a cooperative goal versus a competitive goal. We note that other than the goal and reward the two conditions were identical. Additionally, we examined whether any difference in ToM across cooperation and competition could be attributed to a difference in memory or executive functioning, two cognitive processes known to contribute to ToM. Based on prior testing, we decided to focus solely on 4-year-olds. The study was pre-registered: http://aspredicted.org/blind.php?x=xn73hw.

Methods

Participants

Participants were recruited from public parks in Boston, MA, USA. We had a predetermined goal of 120 participants (60 participants per condition; roughly 30 per gender per condition). Of the 146 participants that were recruited for Study 1, 26 were excluded (breakdown of sample size per cell and detailed information on exclusion criteria are reported in online repository). The final sample thus consisted of 120 4-year-olds, 66 of which were female. Participants were assigned to either the *Cooperation* or the *Competition* condition: mean age did not differ across condition ($M_{\text{Cooperation}} = 4.59$, $M_{\text{Competition}} = 4.56$; $t(116.84) = 0.710$, $p = 0.479$). A legal guardian provided informed consent for all children. This study was approved by the Boston College Institutional Review Board.

Procedure

Participants completed two tasks (scripts provided in online repository): the first task was a two-person game involving stickers, where the goal of the game was to get as many stickers as possible (Fig. 1). A second experimenter, hereafter referred to as E2, was present along with the child during the game instructions. Each participant was assigned to either the *Competition* condition or the *Cooperation* condition. In the *Competition* condition, participants were instructed to hide a sticker in one of two cups while 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 sticker is. E2 would then make a guess as to the sticker's location based solely on the participant's response; although we did not explicitly tell the child, E2 always chose the child's actual selection. In this condition, only one person could win stickers at a time: if E2 guessed correctly, she kept the sticker, but if she guessed incorrectly, the participant got to keep the sticker. In the *Cooperation* condition, the participant was instructed to hide two stickers, with both stickers going in the same cup. In this 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. Crucially, in order to succeed on either task, the participant would have to plant a false belief in E2's mind, so that E2 would subsequently guess incorrectly. We note that we did not use the terms 'deception' or 'lying' in our instructions to participants.

Importantly, participants also responded to two comprehension check questions: (1) If E2 guesses right, who gets to keep it/them? (2) If E2 guesses wrong, who gets to keep it/them? If a question was answered incorrectly, the rules were again described to the participant. If a participant did not report a basic understanding of the game rules after three rounds of explanation, they were excluded from the sample (this affected 3 participants); we had settled on three rounds prior to testing, although we acknowledge we did not specify this number in the pre-registration. Each participant played four rounds of the game; thus, the participant had the opportunity to win up to 4 stickers in either condition. After the game, we asked participants one memory question about the last trial, to check whether the two conditions differed in terms of participants' short-term memory: "Which was the last cup that E2 pointed to?"

After the memory question, participants completed a separate test of inhibition or executive functioning with a variant of the Day-Night task (Gerstadt et al., 1994), which resembles a Stroop task for children. Importantly, we designed the Day-Night task to include a cooperative and competitive condition. The aim of including the Day-Night task was to determine whether any difference found in the Stickers task could be attributed to a difference in executive functioning rather than ToM per se. A condition difference in the Stickers task but not the Day-Night task would be evidence that ToM, but not executive functioning, is sensitive to the difference between cooperative and competitive contexts. The task consisted of 16 trials, in which participants were told to say “Day” in response to an image of nighttime and “Night” in response to an image of daytime. We adapted this task by creating a competitive or cooperative version of it: in the Competition condition, children were told that, if they got more trials correct than E2 (who had played the game before), they would receive a sheet of four stickers (matching the number of stickers a participant could potentially win in the Stickers task). Otherwise, E2 would win a sheet of 4 stickers. Meanwhile, in the Cooperation condition, children were instructed that, if they got more trials correct than E2, they would win a sheet of four stickers for the participant and a sheet of four stickers for E2; otherwise, no one would win anything. These versions mimicked the structure of the conditions in the Stickers task.

Children’s assigned condition remained constant across both tasks; for instance, a child in the Cooperation condition would complete the cooperative version of the Stickers task and the cooperative version of the Day-Night task, whereas a child in the Competition condition would complete the competitive version of each. Task order was counterbalanced across participants.

[insert Fig. 1 here]

Fig 1. Schematic of the Stickers task. In Study 1, 4-year-olds completed the Stickers task with a second experimenter (E2). In Studies 2a and 2b, adults completed an online version of the task.

Analyses

Analyses were conducted in R (version 3.3.3; R Core Team, 2015). In the pre-registration, we stated that we would: (1) perform analyses by creating generalized linear models and performing likelihood ratio tests to compare models with and without predictors of interest, and (2) analyze performance on the Stickers task with generalized linear models. 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 (continuous), and Gender (male or female). We note that the analyses presented here deviated from what was stated in the pre-registration in one main way: we also included the interaction between Condition and Age in our full model, as this is a common interaction tested in the developmental psychology literature. This interaction term was also included in models with response to the memory question (binary: correct or incorrect) and performance on the Day-Night task (proportion of questions correct) as dependent outcomes.

All other analyses reported here are exploratory and were not described in the pre-registration. In one analysis, we compared the full model with a null model that did not include the predictors of interest (Condition and the interaction between Condition and Age) but retained our control predictors (Age, Gender) to test whether our predictors of interest combined had an impact on the proportion of stickers won. This analysis was suggested to us as an omnibus test safeguarding against Type I errors (Forstmeier & Schielzeth, 2011). Other analyses examine performance across trials and number of switches across Condition.

Results

For the Stickers task, our full model containing our predictors of interest (Condition and the interaction between Condition and Age) explained significantly more variance in our response term (proportion of stickers won) than the null model ($\chi^2(3) = 6.982$, $p = 0.03$).

We assessed the significance of our predictors of interest using likelihood ratio tests. There was a significant interaction between Condition and Age ($\chi^2(1) = 5.884$, $p = 0.015$; Fig. 2). That is, younger 4-year-olds won more stickers in the *Cooperation* condition than in the *Competition* condition, whereas there was no difference between conditions in older 4-year-olds. A different way of analyzing the same data (i.e., with a generalized linear mixed model using a binary response variable [winning a sticker versus not] and including participant and trial as random effects) revealed the same two-way interaction ($\chi^2(1) = 6.78$, $p = 0.009$).

We also analyzed the data at the level of individual trials in an exploratory fashion, for three reasons. First, because participants did not initially know that E2 would always guess the sticker's location based solely on the participant's direction, we might see a different pattern for the first trial versus later trials. With more trials, participants might learn how E2 consistently behaved and start responding "correctly". If so, the interaction between condition and age (better performance for cooperation than competition in younger vs. older 4-year-olds) might emerge only for later trials. Second, if participants are indeed learning how to win stickers, another consideration is whether they are learning at different rates, across trials, for cooperation and competition. That is, better performance for cooperation than competition for younger 4-year-olds may be due to younger children learning faster in the case of cooperation. If so, we may see evidence of this by looking at the trial-wise data.

In analyses with Trial as a predictor, there was an effect of Trial ($\chi^2(3) = 14.907$, $p = 0.002$), suggesting that performance improved across trials. Differences across Trial did not vary by Condition ($\chi^2(3) = 0.313$, $p = 0.96$), which suggests that children learned to win stickers at similar rates across conditions, and thus enhanced performance for younger 4-year-olds for cooperation cannot be explained by a difference in the learning rate. We still find an interaction between Condition and Age ($\chi^2(1) = 6.947$, $p = 0.008$). When looking at individual trials, we find the interaction between Condition and Age on trials 2, 3, and 4 (trial 2: $\chi^2(1) = 4.53$, $p = 0.03$; trial 3: $\chi^2(1) = 7.07$, $p = 0.008$; trial 4: $\chi^2(1) = 3.765$, $p = 0.05$), but not on trial 1 (trial 1: $\chi^2(1) = 0.467$, $p = 0.495$) (Figure S1); that is, younger 4-year-olds consistently do better in the *Cooperation* condition than the *Competition* condition beginning on trial 2. No interaction on the first trial but robust interaction on subsequent trials suggest that children display a consistent pattern once they understand how E2 will respond (i.e., E2 will guess the cup that they point to).

We also tested the idea that participants may have used different strategies when playing the game. That is, participants in the *Cooperation* condition may have played consistently (i.e., hiding the stickers in the same cup across trials and pointing to the opposite cup when E2 asked where the stickers were) so that they could make themselves more predictable to E2. On the other hand, participants in the *Competition* condition may have responded in in a more unpredictable manner as an attempt to confuse E2 by hiding stickers in different cups across trials and pointing to different cups when E2 asked where the sticker was. We tested this idea by testing whether the number of switches participants made differed across cooperation and competition: a Welch two-sample *t*-test revealed no significant difference in number of switches across Condition ($t(112.23) = 1.785$, $p = 0.08$).

To examine whether the interaction between Condition and Age found for ToM can be explained by performance on the memory question or performance on the executive functioning task rather than ToM, in particular, we examined whether the same interaction between Condition and Age emerged for both the memory question and the Day-Night task. Analysis of the memory question at the end of the Stickers game revealed no significant interaction between Condition and Age ($\chi^2(1) = 1.66$, $p = 0.20$) and no significant effect of Condition ($\chi^2(1) = 1.83$, *p*

= 0.18). Similarly, analysis of the Day-Night task revealed no significant interaction between Condition and Age ($\chi^2(1) = 0.473$, $p = 0.49$) and no significant effect of Condition ($\chi^2(1) = 2.57$, $p = 0.11$) (Fig. 2). Because neither memory nor response inhibition differed across cooperative and competitive contexts, this suggests that the advantage for cooperation observed for younger 4-year-olds in the Stickers task is specific to ToM.

[insert Fig. 2 here]

Fig. 2. Children's performance in Study 1. Proportion of stickers won in the Stickers task (left) and proportion of correct trials on the Day-Night task (right), by Age and Condition. Error bars denote 95% CIs.

Study 2

In Studies 2a and 2b, we used the same Stickers task to examine whether adults' ToM deployment differed across cooperation and competition. Study 2a had a between-participants design, while Study 2b had a within-participants design. In addition to testing the main question, Study 2a examined the extent to which the number of stickers that could potentially be won could affect participants' responses. Meanwhile, Study 2b examined whether differences in participants' performance on the Stickers task across cooperation and competition could be explained by differences in participants' cognitive effort, game preference, person preference, or subjective responses to deceiving an interaction partner across contexts. In both studies, participants were monetarily compensated for completing the study, with no incentivized bonus related to task performance. While physical stickers were not distributed, participants saw images of stickers. Participants were told they were playing a game with another person, though in reality they were not. We note that in neither of the studies did we use the terms 'deception' or 'lying' in instructions to participants. Study 2a was an exploratory study, and Study 2b was a conceptual replication of Study 2a; the pre-registration for Study 2b can be found here: http://aspredicted.org/blind.php?x=zg4z5y. Both studies were approved by the Boston College Institutional Review Board.

Study 2a

Methods

Participants

We recruited adults 18 years and older using Amazon Mechanical Turk. We aimed for 60 participants per cell. In total, 237 participants completed the task. We excluded participants who failed comprehension checks; the final sample consisted of 194 participants (91 females; M_{age} = 35.96, SD_{age} = 10.70, age range: 19-76).

Procedure

Participants completed the Stickers task online (Fig. 1). We used a 2 x 2 between-participants design: we varied the social context (cooperation vs. competition) and the number of total stickers a participant could win after four trials (4 stickers versus 8 stickers). Sticker number was varied to test whether the same pattern of results would emerge for different reward amounts. Participants were randomly assigned to one of the four conditions. Upon receiving consent, we presented participants with instructions for their variant of the Stickers game. On the same page as the instructions, we asked participants one comprehension question (If the other player guesses correctly and finds the stickers, who gets to keep them?). This provided participants the opportunity to review the instructions if they felt they could not answer the question correctly. On a separate page, we asked participants two questions: one was the same as the one above and the other was: “If the other player guesses incorrectly and doesn't find the stickers, who gets to keep them?”. This provided us with a way of assessing whether participants knew the rules of the game without referring back to the instructions. Participants then completed four trials of the Stickers game and answered a demographics questionnaire.

Analyses

Analyses were conducted in R (version 3.3.3; R Core Team, 2015). We analyzed performance on the Stickers task using generalized linear models. We were primarily interested in whether the proportion of stickers won depended on condition as well as the total number of possible stickers won. To mirror the analyses of Study 1, we included Age and Gender as predictors, though we did not have specific *a priori* hypotheses about these variables. Our full model included the following predictor variables: Condition (cooperation or competition), Total number of possible stickers won (4 or 8), Age (continuous), and Gender (male, female, or other). We also examined the two-way interaction between Condition and Age, given the significance of this interaction in children. 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.

As in Study 1, we created a null model that did not include predictors of interest (Condition and the interaction between Condition and Age) but retained control predictors (Total number of possible stickers won, Age, Gender). We compared the full and null models to test whether our predictors of interest combined have an impact on the proportion of stickers won.

Results

Our full model containing our predictors of interest (Condition and the interaction between Condition and Age) explained significantly more variance in our response term (proportion of stickers won) than the null model ($\chi^2(3) = 47.676$, $p < 0.001$). Unlike the analyses in Study 1, the analyses for this study did not reveal a significant interaction between Condition (cooperation or competition) and Age ($\chi^2(1) = 0.142$, $p = 0.71$). Importantly, however, we found an effect of Condition ($\chi^2(1) = 47.534$, $p < 0.001$), with no effect of Number of possible stickers won ($\chi^2(1) = 0.05$, p =0.82). (Fig. 3a). That is, adults won a greater proportion of stickers in the *Cooperation* condition than in the *Competition* condition, regardless of whether they were given the opportunity to win up to a total of 4 stickers or 8 stickers. Analyzing the same data with a generalized linear mixed model using a binary response variable (winning a sticker versus not) and including participant and trial as random effects revealed the same pattern of results (2-way interaction: $\chi^2(1) = 0.782$, $p = 0.38$; main effect of Condition: $\chi^2(1) = 24.51$, $p < 0.001$; main effect of Number of possible stickers won: $\chi^2(1) = 0.034$, $p = 0.85$).

As in Study 1, we tested whether this pattern emerged for each trial. For example, it could be the case that participants did not initially differ in their performance across contexts but learned to differentiate between contexts over time. In an analysis where the model only included Condition, Trial, and the interaction between Condition and Trial as predictors, we found an effect of Trial ($\chi^2(3) = 90.687$, $p < 0.001$) but no interaction between Trial and Condition ($\chi^2(3) = 3.344$, $p = 0.34$), suggesting that, while participants' performance improved over time, the rate of improvement did not differ across context. Unlike Study 1, there was no interaction between Condition and Age on any trial; instead, there was an effect of Condition for each trial, marginal for trial 1 and significant for trials 2, 3, and 4 (trial 1: $\chi^2(1) = 3.711$, $p = 0.054$; trial 2: $\chi^2(1) = 9.235$, $p = 0.002$; trial 3: $\chi^2(1) = 9.874$, $p = 0.002$; trial 4: $\chi^2(1) = 15.772$, $p < 0.001$). The weaker effect of Condition on Trial 1 but robust effect on subsequent trials demonstrate that once adults learn how their partner will respond to their pointing (e.g., how the game works), they perform better in the *Cooperation* condition than in the *Competition* condition.

As in Study 1, we also tested the idea that participants may have used different strategies when playing the game. That is, participants in the *Cooperation* condition may have played consistently so that they could make themselves more predictable to E2, while participants in the *Competition* condition may have responded in in a more unpredictable manner as an attempt to confuse E2. We tested whether the number of switches participants made differed across cooperation and competition: as in Study 1, a Welch two-sample *t*-test revealed no significant difference in number of switches across Condition ($t(165.36) = 0.11$, $p = 0.91$).

Study 2b

Participants

We recruited adults 18 years and older using Amazon Mechanical Turk. We determined our sample size with a power analysis using the R package 'simr'. With alpha = 0.05 and power = 0.80, the projected sample size needed for a slope of 0.6 (a more conservative estimate than what was observed in Study 2a) is approximately 190. In total, 190 participants completed the task. We excluded participants who failed comprehension checks; the final sample consisted of 174 participants (69 females; $M_{age} = 34.83$, $SD_{age} = 10.18$, age range: 19-64).

Procedure

Study 2b differed from Study 2a in four ways. (1) While Study 2a employed a between-participants design, Study 2b employed a within-participants design. The within-participants design allowed us to examine people's preferences for one game over the other. Participants completed both *Cooperation* and *Competition* conditions of the Stickers task online, presented in a counterbalanced order across participants. Upon receiving consent, participants read instructions for the first variant of the Stickers task to which they were assigned (either the cooperative version or the competitive variant); the instructions were the same as in Study 1a. When participants finished all four trials of the first game, they read instructions for the second task (the condition that they had not yet experienced) and completed four trials of that task. Participants were told that they were playing the second game with a different person. (2) Given that adults' performance on the Stickers task did not depend on whether they could win 4 versus 8 stickers, in this study we presented only the version in which participants could win 4 stickers. (3) We changed the procedure for presenting comprehension check questions to further improve comprehension. Whereas in Study 2a comprehension questions were presented on a separate page from the instructions, in this study, comprehension questions were presented on the same page as the instructions; moreover, we provided participants two chances to get each of the two comprehension questions correct. (4) We introduced new questions at the end of the Stickers task. Participants answered questions about game preference (Which game would you rather play again? Game 1 or Game 2), person preference (Which person would you rather play a different game with? Person from Game 1 or Person from Game 2), strategies for playing each game, feelings about the action they needed to take in order to win a sticker (e.g., deception; In Game X, in order to successfully win a sticker, you had to click on the empty cup when responding to the other player's question. How did you feel about performing that action? Rating was made on a scale from 1 [Extremely good] to 7 [Extremely bad]), and demographic information.

Analyses

Analyses were conducted in R (version 3.3.3; R Core Team, 2015). In the pre-registration, we stated that we would analyze the data in the Stickers task with generalized linear mixed models. We stated that the full model would include whether the participant is able to win a sticker (binary: yes or no) as the dependent measure, Condition, Age, and Gender as predictor variables, and participant and trial as random effects. We also stated that we would test our predictors of interest with likelihood ratio tests, comparing models with and without our predictors of interest. We analyzed performance on the Stickers task using generalized linear mixed models. We were primarily interested in whether the ability to win stickers differed across condition. Our full model included the following predictor variables: Condition (cooperation or competition), Age (continuous), Gender (male, female, or other), Order (first or second), and the interaction between Condition and Age. Though the latter two predictors were not in the pre-registration, we included Order to examine whether the order in which conditions were presented influenced task performance, and the two-way interaction between Condition and Age given the results of Study 1 (though we note we did not have any *a priori* hypothesis regarding the interaction between Condition and Age for adults). We included participant and trial as random effects. For each model, we assessed the importance of our predictors of interest by performing likelihood ratio tests and examining whether the model including a given term provided a significantly better fit to the data than the model without that term. Responses to preference questions were analyzed with chi-square tests. Likert ratings for the feeling question was analyzed with the Wilcoxon signed-rank test.

Similar to the previous two studies, we performed other analyses not specified in the pre-registration. Mainly, for the Stickers task, we created a null model that did not include our predictors of interest (Condition and the interaction between Condition and Age) but retained our control predictors (Age, Gender, Order). We compared the full model with the null model to test if our predictors of interest combined had an impact on whether a sticker was won. Other analyses described here were exploratory.

Results

Our full model containing our predictors of interest (Condition and the interaction between Condition and Age) explained significantly more variance in our response term (whether a sticker was won) than the null model ($\chi^2(2) = 49.727$, $p < 0.001$). As in Study 2a, our analyses did not reveal a significant interaction between Condition and Age ($\chi^2(1) = 1.735$, $p = 0.19$). However, there was a significant effect of Order ($\chi^2(1) = 73.102$, $p < 0.001$): participants performed better on the second variant than the first variant. Most importantly, as in Study 2a, there was again an effect of Condition ($\chi^2(1) = 47.992$, $p < 0.001$): adults were better at winning stickers in the *Cooperation* condition than in the *Competition* condition (Fig. 3a).

As we did in Study 1 and Study 2a, we tested for this pattern across trials in an exploratory fashion. In an analysis where the model included only Condition, Trial, and the interaction between Condition and Trial as predictors, we found an effect of Trial ($\chi^2(3) = 90.09$, $p < 0.001$), suggesting that performance improved across trials, and an effect of Condition ($\chi^2(1) = 44.612$, $p < 0.001$), consistent with Study 2a. Notably, the effect of Condition was found for each trial (trial 1: $\chi^2(1) = 6.355$, $p = 0.012$; trial 2: $\chi^2(1) = 10.219$, $p = 0.001$; trial 3: $\chi^2(1) = 16.502$, $p < 0.001$; trial 4: $\chi^2(1) = 22.376$, $p < 0.001$), with no interaction between Condition and Age (trial 1: $\chi^2(1) = 0.001$, $p = 0.98$; trial 2: $\chi^2(1) = 1.062$, $p = 0.303$; trial 3: $\chi^2(1) = 0.248$, $p = 0.62$; trial 4: $\chi^2(1) = 2.45$, $p = 0.12$). These results suggest once again that adults perform better in the Stickers task in the *Cooperation* condition than in the *Competition* condition.

Again, as in Study 1 and 2a, we tested the idea that participants may have used different strategies when playing the game. We tested whether the number of switches participants made differed across cooperation and competition: a Welch two-sample *t*-test revealed a significant difference in number of switches across Condition ($t(345.07) = 2.398$, $p = 0.02$): participants switched their responses more in the *Competition* condition than in the *Cooperation* condition. Further investigation into this result showed that this difference was not found for the first task variant presented ($t(171.26) = 0.543$, $p = 0.59$) but instead found for the second task variant presented ($t(171.77) = 2.817$, $p = 0.005$). That is, when we compare number of switches between the *Competition* condition for participants receiving the competition variant first and the *Cooperation* condition for participants receiving the cooperation variant first, we find no difference. We only see a difference between condition for the second variant presented to participants. This result suggests that the strategy used does change across conditions, but after exposure to a different variant of the game.

Next, we examined other factors that could potentially account for the difference in performance across contexts. While response time differed across Order (i.e., participants took longer to respond on the first task than the second task; $\chi^2(1) = 40.860$, $p < 0.001$), response time did not differ across Condition ($\chi^2(1) = 0.126$, $p = 0.723$). This result suggests that a difference in performance across cooperation and competition is unlikely due to a difference in task difficulty or cognitive effort, at least as indexed by reaction time. We also assessed whether people showed a preference for the cooperative game or the competitive game: people showed a preference for the cooperative game (N=120 for cooperation versus N=54 for competition; $\chi^2(1) = 25.034$, $p <$ 0.001; Fig. 3b). People also showed a preference for the person with whom they played the cooperative game (N=105 for cooperation versus N=69 for competition; $\chi^2(1) = 7.448$, $p = 0.006$; Fig 3b). Finally, people rated their feelings, from a scale of 1 (Extremely good) to 7 (Extremely bad), differently across the two conditions ($V = 2977.5$, $p < 0.001$; Fig. 3c): people rated their feelings more positively in the *Cooperation* condition than in the *Competition* condition. However, entering response time, game preference, person preference, and feelings as predictors in the model did not affect the pattern of results: again, we find a significant effect of Condition ($\chi^2(1) = 33.732$, $p < 0.001$).

[insert Fig. 3 here]

Fig 3. Adults' responses on the Stickers task in Studies 2a and 2b. (A) Performance on the Stickers task. (B) Responses to preference questions (Study 2b). (C) Ratings regarding feelings about deceiving their partner in the Stickers task (Study 2b). Note that we did not use the word deception in the question (question paraphrased here for visualization purposes; see methods for exact question asked). Error bars denote 95% CI.

General Discussion

Three studies demonstrate differences in people's abilities to plant false beliefs in others in order to achieve cooperative versus competitive goals. While some researchers have examined the influences of contextual factors such as social status on ToM performance (Kraus, Piff, Mendoza-Denton, Rheinschmidt, & Keltner, 2012; Rizzo & Killen, 2018), here, we examine how people may differ in their deployment of ToM across different social contexts. By considering the impact of socio-moral context on ToM in children and adults, the present work treats moral psychology as a hub connecting social cognition and developmental psychology to theories in evolutionary psychology and intergroup cognition. Specifically, the present work reveals differences in how people consider other minds (e.g., friend or foe) in development and in adulthood. Study 1 reveals an age-dependent difference in performance across cooperation and competition. We find that, younger 4-year-olds (but not older 4-year-olds) are better able to plant false beliefs in another's mind to achieve a cooperative goal versus a competitive goal. Importantly, this difference between cooperation and competition appears specific to ToM and does not emerge for executive functioning. Meanwhile, Studies 2a and 2b reveal that adults are more successful at planting false beliefs in others in cooperative versus competitive contexts. We highlight that, even though older 4-year-olds did not show a difference in ToM between cooperation and competition, any time a difference did emerge—in children (i.e., younger 4-

year-olds) or in adults—the difference was in the direction of a cooperative advantage. This directional pattern is especially noteworthy given that success on the Stickers task required deceiving another person, and typically deception is more closely linked to achieving competitive or otherwise selfish goals. Indeed, one might predict that people would be *better* at planting false beliefs to achieve a *competitive* goal than a cooperative goal, but the present work provides initial evidence to the contrary.

Why might cooperation boost ToM performance among younger children engaging in actual interactions? One possibility is that younger children may have a propensity to help those around them. Indeed, infants as young as 14 months of age readily help others to achieve their goals (e.g., helping others get out-of-reach objects) (Warneken & Tomasello, 2006; Warneken & Tomasello, 2007). Not only do infants engage in instrumental helping, but also they engage in emotional helping (e.g., comforting others in distress) (Johnson, 1982; Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992). However, young children are not indiscriminate helpers: children as young as 3 years old have been shown to provide help in a selective manner (e.g., children will help pick up a functional but not a dysfunctional object for another person; Martin & Olson, 2013). Perhaps younger children may be more likely to consider the minds of others if doing so improves their ability to actively and *effectively* cooperate with and help others. Human societies display an immense capacity for large-scale cooperation and altruistic social preferences that is uncommon in other species (Warneken & Tomasello, 2009); this capacity may boost early false belief understanding (Matsui & Miura, 2008). Our work supports the idea that ToM is developed to support large-scale cooperation found in human societies. But we also contribute to this body of literature by revealing nuances associated with this claim: we demonstrate that a ToM advantage for cooperation is present for younger preschoolers, that it emerges for interactions with others, and that it is specific to ToM and not a result of other cognitive processes like executive functioning. Therefore, at a time when children are less fluent in explicit ToM, a time when condition differences might be expected, we indeed see a difference in the direction of enhanced performance for cooperation versus competition.

To our knowledge, our work is the first to document a developmental shift in ToM across cooperative and competitive contexts during the preschool years, from enhanced ToM for cooperation than competition in younger preschoolers to a null effect for older preschoolers, and another shift to enhanced ToM for cooperation than competition in adulthood. This finding raises important questions about the mechanisms underlying these changes: (1) Do younger and older preschoolers focus on different aspects of the task? Perhaps older preschoolers focus more on attaining the reward rather than the cooperative or competitive nature of the task, leading to different behavioral performance. However, in pilot testing the task on children using different reward amounts (4 versus 8 stickers), we did not find reward amount to affect children's performance on the Stickers task. (2) Are younger children's tendency to apply their understanding of false beliefs in cooperative contexts qualitatively different from adults' tendency to do the same? That is, in younger children, this bias to engage in cooperative ToM might reflect a natural predilection, whereas in adults the same tendency might reflect enhanced motivation to plant false beliefs to achieve cooperative goals, given adults' preference for and positive feelings toward cooperation. We have some evidence to support that this isn't the case for adults: after accounting for adults' preference for cooperation, preference for the cooperative partner, and positive feelings about cooperation, adults' enhanced performance for cooperation persists. We view our studies as providing a basis for future investigations on mechanisms supporting young children's and adults' ToM bias toward cooperation.

Addressing alternative explanations

While prior work has measured the extent to which adults process mental state information across cooperation and competition (Tsoi, Dungan, Waytz, & Young, 2016), the current work focuses on how *effective* people are at using or manipulating mental states in these two social contexts. Indeed, one key contribution of the present work is that the Stickers task allows us to explore variation in ToM success, measured objectively, in adults. Standard false belief tasks are not typically used to measure variation in ToM success, given that adult performance is at ceiling. By contrast, the Stickers task allows for quantification of *both* children's and adults' ToM performance in the same experimental context; moreover, in using this task, we are examining not just false belief understanding, but also whether people can deploy this understanding to actively influence others' beliefs.

However, one key question that arises is whether the Stickers task adequately captures ToM. We note that our task shares many similarities with other tasks measuring ToM that also utilize deception (e.g., Chandler & Hala, 1994; Sher, Koenig, & Rustichini, 2014; Sullivan & Winner, 1993), with the caveat that we created tightly-controlled cooperative and competitive versions of our task. In children, we tested for condition differences in executive functioning or memory, but we did not find the same Condition by Age interaction that we found for the Stickers task, leading us to rule out the possibility that the Stickers task could be instead tapping into executive functioning or memory.

We also argue against several different interpretations of the task. One interpretation is that the task may be capturing differences in strategies. Instead of employing ToM, participants may act more predictably in cooperative contexts (Vesper, van der Wel, Knoblich, & Sebanz, 2011) and less predictably in competitive contexts (Ybarra et al., 2010). Indeed, one might imagine that participants competing with another would switch their responses frequently to confuse the other player. However, in our studies with between-participants designs, we do not see a significant difference in number of switches across social contexts.

A different but related concern is that people may be using higher-order ToM or recursive thinking during competition but not during cooperation, and that this recursion, at least in the context of our task, leads people to do poorly when competing with another. However, if this were the case, we would, again, see people switching their responses more during competition than during cooperation, but this is not what we find. Moreover, prior work has revealed that children's ability to engage in strategic reasoning, including recursive theory of mind, appears around 7 years of age (Sher et al., 2014), which suggests that this explanation may apply less well for understanding the developmental shift we find for preschoolers.

We argue against a third interpretation of the task, namely that participants may not be representing their partner's beliefs at all but instead are relying on a rule-based approach (e.g., noting the association between outcomes and their pointing to specific cups). First, we note that the Day-Night task, which requires a rule-based approach for optimal performance, does not elicit differential performance across contexts. Second, we have some evidence to suggest that adults are indeed representing their partner's belief. That is, despite not using the term 'deception' or any other related term throughout the task, many adults provide explicit mentions of deception in their responses to a post-task question about game strategy (Study 2a: out of 161 participants who responded to using a strategy, 62 of these participants explicitly and spontaneously mentioned deception or other synonyms).

A different line of inquiry taps into understanding why younger 4-year-olds may have performed worse during competition than cooperation, and whether reasons unrelated to ToM could explain the effect. One possibility is that younger 4-year-olds tried to be nice to E2 even when competing with E2, thereby letting E2 win more. However, past research in the domain of fairness has shown that children at this age are relatively selfish, in that they tend to favor themselves versus others when it comes to receiving rewards such as stickers (McAuliffe, Blake, Steinbeis, & Warneken, 2017), making this explanation less probable. A related possibility is that younger children tend to be more trusting of others, especially adults, which may impact their ability to successfully engage in competitive behaviors against another. Indeed, some work has found that children younger than 5- or 6-years old are generally trusting of others, even those that provide deceptive testimonies to them (Jaswal, Croft, Setia, & Cole, 2010). If this were the case, we would expect to see worse performance for competition than cooperation across age; however, that's not the pattern that we find. Moreover, other work suggests that children around 5-years old display different levels of trust in others depending on whether they share aligned (cooperative) versus conflicting (competitive) interests (Reyes-Jaquez & Echols, 2015), making it less likely for trust to be a main predictor of our effect. Other possibilities, such as being too intimidated to overtly mislead the experimenter in order to win stickers, or having expectations of turn-taking (that by letting E2 win in the current trial E2 will let the participant win in future trials), also seem unlikely given that children show consistent improvement across trials, with no differences in improvement across cooperation and competition.

Conclusion

At the broadest level, we demonstrate differences in ToM deployment across social contexts: we find evidence of more successful ToM deployment for cooperation than competition in both younger 4-year-olds and adults. While it is currently unclear whether the same mechanisms support this ToM advantage for cooperation in young children and adults, we have reason to believe that response time, preference for one context over the other, preference for one game partner over the other, and feelings about deceiving a partner cannot account for enhanced ToM for cooperation. Together, these results provide preliminary evidence that children initially display a natural tendency to deploy ToM for cooperation over competition, and that this pattern is exhibited in adulthood as well. By housing our work in the moral domain and linking our work to perspectives in social cognition, developmental psychology, intergroup cognition, and evolutionary psychology, we are able to reveal new insights into how people navigate social and moral contexts. Specifically, this approach affords four key contributions: (1) with respect to social cognition, we demonstrate that a key social cognitive process—ToM—is deployed differently across social contexts, (2) with respect to developmental psychology, we document a cooperative ToM advantage in young 4-year-olds, (3) with respect to social psychology or intergroup cognition, we show that socio-moral context (e.g., friend or foe) affects ToM, and (4) with respect to evolutionary psychology, we provide initial evidence consistent with the idea that ToM has developed to support large-scale cooperation found in human societies.

Acknowledgments

We thank the following for their assistance with data collection: [Study 1] Catherine Kim, Elizabeth Kroll, Joshua Hirschfeld-Kroen, Kayla Soares, Michael Manalili, and [preliminary testing] Sophie Cooper. We also thank members of the Boston College Morality Lab and anonymous reviewers for comments on earlier versions of this manuscript, the staff of the Living Lab and Science World at TELUS World of Science, and participating families. This research was supported by the Social Sciences and Humanities Research Council of Canada (A.S.B.: # 435-2013-0286) and the National Science Foundation (L.Y. and A.W.: # 1627157; graduate fellowship to L.T.: # 1258923).

References

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

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.

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

Carlson, Stephanie M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development*, *69*(3), 672–691.

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

Chandler, M., & Hala, S. (1994). The role of personal involvement in the assessment of early false belief skills. *Children's Early Understanding of Mind: Origins and Development*, 403–425.

Ding, X. P., Heyman, G. D., Fu, G., Zhu, B., & Lee, K. (2018). Young children discover how to deceive in 10 days: A microgenetic study. *Developmental Science*, *21*(3), e12566. https://doi.org/10.1111/desc.12566

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology*, *65*(1), 47–55. https://doi.org/10.1007/s00265-010-1038-5

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 312–7 years old on a stroop- like day-night test. *Cognition*, *53*(2), 129–153. https://doi.org/10.1016/0010-0277(94)90068-X

Gordon, A. C. L., & Olson, D. R. (1998). The Relation between Acquisition of a Theory of Mind and the Capacity to Hold in Mind. *Journal of Experimental Child Psychology*, *68*(1), 70–83. https://doi.org/10.1006/jecp.1997.2423

Grueneisen, S., Duguid, S., Saur, H., & Tomasello, M. (2017). Children, chimpanzees, and bonobos adjust the visibility of their actions for cooperators and competitors. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-08435-7

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

Hamlin, J. K., & Baron, A. S. (2014). Agency attribution in infancy: Evidence for a negativity bias. *PLoS ONE*, *9*(5), e96112. https://doi.org/10.1371/journal.pone.0096112

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

Jaswal, V. K., Croft, A. C., Setia, A. R., & Cole, C. A. (2010). Young Children Have a Specific, Highly Robust Bias to Trust Testimony. *Psychological Science*, *21*(10), 1541–1547. https://doi.org/10.1177/0956797610383438

Johnson, D. B. (1982). Altruistic behavior and the development of the self in infants. *Merrill-Palmer Quarterly*, *28*(3), 379–388.

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

Knobe, J. (n.d.). *Intentional action and side effects in ordinary language*. 5.

Kraus, M. W., Piff, P. K., Mendoza-Denton, R., Rheinschmidt, M. L., & Keltner, D. (2012). Social class, solipsism, and contextualism: How the rich are different from the poor. *Psychological Review*, *119*(3), 546–572. https://doi.org/10.1037/a0028756

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

Leslie, A. M., Knobe, J., & Cohen, A. (2006). Acting intentionally and the side-effect effect: Theory of mind and moral judgment. *Psychological Science*, *17*(5), 421–427. https://doi.org/10.1111/j.1467-9280.2006.01722.x

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

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

Martin, A., & Olson, K. R. (2013). When kids know better: Paternalistic helping in 3-year-old children. *Developmental Psychology*, *49*(11), 2071–2081. https://doi.org/10.1037/a0031715

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

McLoughlin, N., & Over, H. (2017). Young children are more likely to spontaneously attribute mental states to members of their own group. *Psychological Science*, *28*(10), 1503–1509. https://doi.org/10.1177/0956797617710724

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

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

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

Reyes-Jaquez, B., & Echols, C. H. (2015). Playing by the rules: Self-interest information influences children's trust and trustworthiness in the absence of feedback. *Cognition*, *134*, 140–154. https://doi.org/10.1016/j.cognition.2014.10.002

Rizzo, M. T., & Killen, M. (2018). How social status influences our understanding of others' mental states. *Journal of Experimental Child Psychology*, *169*, 30–41. https://doi.org/10.1016/j.jecp.2017.12.008

Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The 'windows task' as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology*, *9*(2), 331–349. https://doi.org/10.1111/j.2044-835X.1991.tb00881.x

Sher, I., Koenig, M., & Rustichini, A. (2014). Children's strategic theory of mind. *Proceedings of the National Academy of Sciences*, *111*(37), 13307–13312. https://doi.org/10.1073/pnas.1403283111

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.

Sullivan, K., & Winner, E. (1993). Three-year-olds′ understanding of mental states: The influence of trickery. *Journal of Experimental Child Psychology*, *56*(2), 135–148. https://doi.org/10.1006/jecp.1993.1029

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

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., & Young, L. (2018). Moral Reasoning. In J. T. Wixted (Ed.), *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (pp. 1–25). https://doi.org/10.1002/9781119170174.epcn312

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

Vesper, C., van der Wel, R. P. R. D., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: Reduced temporal variability facilitates joint action coordination. *Experimental Brain Research*, *211*(3–4), 517–530. https://doi.org/10.1007/s00221-011-2706-z

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301–1303. https://doi.org/10.1126/science.1121448

Warneken, Felix, & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, *11*(3), 271–294. https://doi.org/10.1111/j.1532-7078.2007.tb00227.x

Warneken, Felix, & Tomasello, M. (2009). The roots of human altruism. *British Journal of Psychology*, *100*(3), 455–471. https://doi.org/10.1348/000712608X379061

Waytz, A., Morewedge, C. K., Epley, N., Monteleone, G., Gao, J.-H., & Cacioppo, J. T. (2010). 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

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.

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

Ybarra, O., Keller, M. C., Chan, E., Garcia, S. M., Sanchez-Burks, J., Morrison, K. R., & Baron, A. S. (2010). Being Unpredictable: Friend or Foe Matters. *Social Psychological and Personality Science*, *1*(3), 259–267. https://doi.org/10.1177/1948550610370214

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, *28*(1), 126–136. https://doi.org/10.1037/0012-1649.28.1.126