UNIVERSITYOF BIRMINGHAM

Individual differences in children's co-representation of self and other in joint action

Milward, Sophie; Kita, Sotaro; Apperly, Ian

Link to publication on Research at Birmingham portal

Individual Differences in Children's Corepresentation of Self and Other in Joint Action

Sophie J. Milward
Max Planck Institute for Evolutionary Anthropology and University of Birmingham

Sotaro Kita
University of Warwick

Ian A. Apperly
University of Birmingham

Previous research has shown that children aged 4–5 years, but not 2–3 years, show adult-like interference from a partner when performing a joint task (Milward, Kita, & Apperly, 2014). This raises questions about the cognitive skills involved in the development of such "corepresentation (CR)" of a partner (Sebanz, Knoblich, & Prinz, 2003). Here, individual differences data from one hundred and thirteen 4- to 5-year-olds showed theory of mind (ToM) and inhibitory control (IC) as predictors of ability to avoid CR interference, suggesting that children with better ToM abilities are more likely to succeed in decoupling self and other representations in a joint task, while better IC is likely to help children avoid interference from a partner's response when selecting their own response on the task.

Many have argued that children are able to participate in interactive, collaborative, and coordinated tasks from an early age, with evidence coming from naturalistic games such as throwing a ball back and forth to one another (Hay, 1979) as well as more controlled, artificial puzzle-solving tasks (Warneken, Chen, & Tomasello, 2006; Warneken, Gräfenhain, & Tomasello, 2012). However, there has been substantial disagreement about whether to interpret these findings as involving high- or low-level cognitive skills (Brownell, 2011). Two recent studies (Milward, Kita, & Apperly, 2014; Saby, Bouquet, & Marshall, 2014) provided a new avenue for investigating early joint action behaviors, by showing that children exhibit adult-like interference effects resulting from corepresentation (CR; Sebanz, Knoblich, & Prinz, 2003) of both one's own and one's partner's task in a joint action. That is, when two individuals play two different roles in a joint action (e.g., one responds to one set of stimuli and the other responds to a complementary set), one's performance can be influenced by what the other person has to do. Such an effect is evidence that participants represent not only their own task goal but also the joint action partner's task goal. In many joint action scenarios, this should aid prediction of a partner's actions and so improve the quality of the interaction (e.g., synchrony between actors). However, in some situations, interference can be caused by their representation of their partner's actions, and this effect has been exploited in experimental settings to demonstrate the existence of CR. In order to test whether high-level cognitive skills are involved in CR effects, we carried out an individual difference study, in which we examined whether 4- to 5-year-olds' performance on the CR task in Milward et al. (2014) was associated with their abilities in theory of mind (ToM), inhibitory control (IC), and working memory (WM).

Bratman (1992) provides an influential definition of joint action that requires, among other things, an understanding of joint intentionality. In other words, each participant in a joint task must understand the other's intention to perform certain actions to reach a joint goal, and that both intend for their action plans to mesh in order to achieve

Correspondence concerning this article should be addressed to Sophie J. Milward, Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6,04103 Leipzig, Germany. Electronic mail may be sent to sophie_milward@eva.mpg.de, milwards@ceu.edu.

that mutual goal. Brownell (2011) points out that this causes a problem for developmental psychologists who are studying a population with arguably immature cognitive skills. Accordingly, one must argue either that early behaviors show genuine joint action understanding (as specified by Bratman, 1992) because they do have the prerequisite skills required (Carpenter, 2009) or that one must explain early behaviors through lower level skills because young children lack the resources for such higher level processing (Butterfill, 2011).

Many studies have found evidence for infants' and preschoolers' ability to perform actions that are joint at least in some sense. In her review of this literature, Carpenter (2009) argues that infants have all the prerequisites needed to perform joint action, even by Bratman's account. Such prerequisites include understanding other's goals and intentions (Behne, Carpenter, Call, & Tomasello, 2005) and understanding of common knowledge (Moll, Richter, Carpenter, & Tomasello, 2008). Behaviors used as evidence for joint action include spontaneous helping (Warneken & Tomasello, 2006: 18-month-olds and chimpanzees), production of declarative pointing in order to reach joint attention (Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004: 12-month-olds), and comprehension of a partner's pointing as related to a joint activity (Liebal, Behne, Carpenter, & Tomasello, 2009: 14-month-olds). However, it is unclear whether any of these behaviors require simultaneous and separate representations of the intentions of self and other, which are necessary for joint intentionality understanding.

Studies attempting specifically to demonstrate joint intentionality understanding have used children's reengagement of a task partner as evidence of such understanding. Most recently, Warneken et al. (2012) found that children were more likely to attempt to reengage a partner who stops participating in a joint task when they were unable rather than unwilling to continue. These authors argue that this is evidence that children at 1–2 years of age are able to perform well on such tasks due to their ability to understand joint intentions. However, these behaviors could also be explained via much more low-level mechanisms (e.g., Brownell, 2011; Butterfill, 2011). For example, children do not need to be able to understand intentions in order to view these two conditions as different. Instead, they could understand that in the "unable" condition there is a reason why the experimenter stops participating (failure to complete the action), whereas in the "unwilling" condition there is not. Thus, it might be more fruitful to attempt to reengage a partner who has failed for

a specific reason (e.g., lacking dexterity) that may be overcome than one who has failed for a reason unavailable to the child. To give an analogy, if my computer fails to perform a task because it does not have the software to do so, I can reengage the computer to rectify this fairly simply by installing the appropriate software. If it fails for no apparent reason, I might just give up. This does not entail consideration of the computer having mental states. Likewise, Butterfill (2011) points out that children could use lower level behavioral goal tracking, rather than intention understanding, in order to perform tasks with others. Thus, there is, as yet, no clear evidence for the involvement of mental state understanding in early joint action behaviors that cannot be explained through lower level mechanisms.

Recent evidence for CR in young children (Milward et al., 2014; Saby et al., 2014) could shed some light on this problem. CR is a mechanism that has been used to explain interference effects between self and other when two individuals engage in a joint action. The original proponents of this explanation (Sebanz et al., 2003) found interference effects on a joint Simon task that were similar to the interference effects found on an individual Simon task. Participants were presented with a hand with an extended index finger either pointing to left or right. The finger wore either a green or red ring, depending on the trial. Participants' task was to press a button on the left of the keyboard for one color and on the right for the other, and the direction in which the hand on the screen pointed was not relevant to their responses. In an individual two-response task, participants were required to respond to both green and red rings with the two buttons. Compatibility effects were found whereby responses were slower when the pointing direction of the finger was incompatible with the location of the correct response button. In a joint version, the task was divided between two actors, so that each was responsible for either green or red rings, and thus for only one response each. The same compatibility effects were also observed in this task, which was surprising given that no such effects were found in an individual, single response task in which a single participant had to respond to a single button alone (without a joint actor).

This original effect was explained as being due to participants in the joint action condition "corepresenting" their task partner and therefore being subject to interference from the representation of the alternative part of the task, which was not represented in an individual, single response version. Thus, this is evidence for the involvement of

simultaneous representation of self and other in adult joint action, which sometimes interfere with one another. This is particularly interesting given that participants were not explicitly asked to work toward a joint goal or to take into consideration the other person's actions or task in any way. The fact that they did this any way suggests a mechanism of joint action involving self–other representation that is more implicit and shows stronger characteristics of automaticity than those implied by the developmental studies reviewed above (Carpenter, 2009). Tasks of this nature could therefore be a promising way to study the processes underlying early joint actions, because they meet some of the criteria for an action to be joint—simultaneous representation of self and other—but do not entail explicit understanding of mutually constrained intentions (Bratman, 1992).

Recently, two studies have indeed found similar effects in children (Milward et al., 2014; Saby et al., 2014). Using different methodologies, each analogous to Sebanz et al.'s (2003) methods, these studies both found CR effects after the age of 4 years. In the task employed by Milward et al. (2014), children were presented with trials in which they saw either a bear or a duck. In one condition (same task), both participants had to press their button when they saw one of the animals (e.g., bear) and inhibit responding to the other (e.g., duck). In the other condition (different task), each participant had to respond to a different animal (e.g., one to the bear and one to the duck). Results showed that performance was worse in the different task condition, suggesting that children had corepresented their partner's task and therefore experienced interference when that task was different from their own. Additionally, Milward et al. (2014) failed to find CR effects in a younger age group aged 2–3 years.

This apparent developmental cutoff point at 4 years suggests CR may be a mechanism of joint action that only develops after this age. If this is the case, it is important to ask what cognitive processes are linked to the onset of CR and its subsequent interference effects. Two obvious potential contributors are ToM (although see Dolk, Hommel, Prinz, & Liepelt, 2013 for a nonsocial explanation of interference effects) and executive abilities that show significant development between 3 and 5 years. Investigation of the former can also speak to the argument surrounding the involvement of representational capacities in joint action in young children.

In relation to ToM, one clear possibility is that developing ToM abilities allow children to construct representations of their task partner's intentions and so engage in CR. In terms of observable effects on the bear–duck task (as employed by Milward et al., 2014), this would mean that developing ToM abilities would lead to representation of the partner's task as well as the child's own, giving rise to the interference effects that are indicative of CR. This possibility would explain why CR effects have only been found at age 4 years, as this corresponds with typical onset of explicit false belief understanding. Thus, children younger than 4 might only represent their own part of the task, completely ignoring and subsequently experiencing no interference from a partner's role. This would fit with the body of research on egocentrism in perspective taking in early childhood (Doherty, 2008; Wimmer & Perner, 1983).

The prediction that the development of ToM abilities enables CR would also fit with findings that CR in adults involves mental state understanding. For example, Humphreys and Bedford (2011) found that patients with lesions associated with ToM deficits did not show spatial compatibility effects in a joint Simon task, whereas they did so in an individual version. These patients showed deficits on measures of false belief understanding, suggesting not only that ToM but specifically belief comprehension is necessary for CR. Other studies looking at mental state understanding and the social Simon effect have focused on intention understanding as a prerequisite for CR. Tsai, Kuo, Hung, and Tzeng (2008) found behavioral and electrophysiological effects of compatibility when participants were told they were performing a Simon task with a human partner but not when told they were playing alongside a computer. Likewise, using a joint Flanker task, Atmaca, Sebanz, and Knoblich (2011) found that participants only showed compatibility effects in a joint task when their partner's actions were intentional rather than when they were controlled by a magnet. Thus, if CR requires explicit mental state understanding, and explicit mental state understanding only develops at around 4–5 years, this would explain why studies have failed to show CR effects below this age.

However, the long-standing idea that developments in ToM may give rise to improvements in executive self-control (e.g., Perner, 1991) predicts the opposite relationship, whereby improving ToM leads to reduced interference from CR of the partner. On this account, whatever the developmental origin of CR effects, one consequence of advances in ToM would be that children are better able to maintain separate, stable representations of "self" and "other" and their respective roles during a joint task. In Perner's (1991) original conception developing ToM provided metacognitive abilities that

enabled self-control through domain-general improvements in executive control. However, an alternative conception arises from recent work suggesting that there may be a social-specific control mechanism involved in decoupling of self and other representations, which has been shown to be distinct from general IC and other processes associated with ToM (Schuwerk et al., 2014).

Recently, several researchers have put forward the idea that self–other distinction may, in fact, be an overarching mechanism in social cognition that links processing in motor, cognitive, and affective domains (Decety & Lamm, 2007; Santiesteban et al., 2012; Sowden & Catmur, 2015; Spengler, von Cramon, & Brass, 2009). For example, Santiesteban et al. (2012) carried out a training study whereby training of imitation inhibition on a finger-tapping (motor domain) task transferred to improvements in visual perspective taking (cognitive domain). They argued that this training improved participants' ability to maintain and manipulate distinct representations of self and other, which is a mechanism that is common across domains in social cognition in general and has been linked to commonly identified "social" brain areas, including medial prefrontal cortex and temporoparietal junction. Of course, this prediction that improving ToM leads to reduced interference from the partner's task once the ability to corepresent is established leaves open the possibility that at an earlier stage in development, developing ToM abilities (specifically relating to construction of representations of self and other rather than resolving conflict between the two) contribute to the onset of CR. Thus, our two predictions about the relationship between CR and ToM are not mutually exclusive across the full course of development, but they do result in opposite predictions for a given age range. The current study aims to distinguish between these two predictions.

By looking at the relationship between CR interference effects and ToM in children who already show evidence of CR, we can investigate first whether there is any relationship with ToM at all, and second what is the direction of this relationship. If CR interference effects increase with developing ToM, then the first of our predictions, that ToM plays a role in establishing CR (i.e., constructing self and other representations), is supported. If CR interference decreases with developing ToM, then the second prediction, that ToM helps children to resolve conflict between self- and other-tagged representations, is supported.

In addition to ToM, another cognitive skill that undergoes significant improvement at around 4–5 years is general IC capacity (Carlson, 2010). Tasks used in both adult and developmental CR research, such as the Social Simon (Sebanz et al., 2003) and Joint Flanker tasks (Atmaca et al., 2011) have made use of IC to demonstrate how representing a partner during a joint task can be detrimental when trying to inhibit a stimulus or stimulus feature on which one's partner is acting. This leads to the reasonable expectation that IC will continue to be a key demand in CR tasks in that CR makes it more difficult to inhibit competing task features. This could be the case whether or not there is also a social-specific component to task performance, and also whether participants represent their partner's task as such (i.e., "my" task vs. "your" task), or whether they simply represent the other part of the task without relating it to another agent (["My"] Task A vs. Task B). When representing a partner, one needs to inhibit the other task role so that they do not interfere with one's own actions or task role. Thus, a greater ability to inhibit the other task role should lead to better performance on a joint task.

In the task employed by Milward et al. (2014), the *same* task should not require inhibiting the other person's task/action (although inhibition is required in general to complete the task), because there is no competition between participants' responses. If anything, inhibiting one's partner should slow responses. In contrast, the *different* task should require inhibition to resist interference from the partner's (different) task. Therefore, it is possible that individual differences in performance on the different task could be modulated, not only by the level to which a person corepresents a partner but also by their capacity for inhibiting the alternative task response option in order to avoid interference. Thus, performance on IC measures should be positively related to performance on a CR task. This should be particularly so for *conflict* IC tasks, given that both involve conflict between two rules rather than a simple delay in response (Carlson & Moses, 2001). This is analogous to the CR task, where there is a conflict between one's own rule and ones partner's rule, even though switching between the two is not explicitly required.

Another *executive* skill that may be involved in CR tasks is WM. Participants must remember their task rule and maintain this while completing the task. If they are to corepresent, they also need to remember the other person's task rule, particularly in versions of the task where participants cannot see each other. This is likely to be highly important in real-world joint action scenarios, which often require adaptation of actions in order to coordinate with a partner. In order to do this, it is necessary to

understand the partner's role in the task. This may also be relevant to experimental designs where tasks with identical versus complementary roles are used, as it is vital to ensure that task types are matched for WM demands. By identifying the role of WM in both conditions in the current study, direct evidence can be obtained as to whether WM has a confounding influence on these tasks.

Set shifting was also considered as a possible measure here, it being generally acknowledged as one of the key components in *executive function* alongside WM and IC (Miyake, Friedman, Emerson, Witzi, & Howerter, 2000). However, this was considered to be less relevant than the latter components, as the go–no-go task employed here does not require participants to act upon two rules which they must switch between. Instead, they must act upon one rule and avoid conflict from an irrelevant rule, as in conflict IC tasks as outlined above.

Each of these cognitive skills has been extensively investigated in terms of individual differences in children's abilities and how these differences relate to other cognitive constructs (Carlson, Moses, & Breton, 2002) as well as real-world outcomes (Blair & Razza, 2007). The question of which capacities might be involved in CR will help us explain the developmental trajectory that we see in early CR tasks. For this reason, the following experiment investigates the relationship between individual differences in WM, IC, ToM, and CR. Specifically, it focuses on which cognitive skills predict individual performance on the computerized corepresentation task from Experiment 3 in Milward et al. (2014). Given that we are interested in abilities related to CR rather than general joint task requirements, the measure of interest is performance on the different task once abilities shared with the same task have been controlled for.

In order to increase validity of the measures used, multiple tasks were administered for each cognitive construct. Tasks were taken from existing studies that have looked at individual differences in executive functions and ToM that have used a similar age range to that in the current study. WM was measured using Backward Digit Span and Counting and Labeling tasks, based on Carlson et al. (2002). IC was measured using Day–Night (Gerstadt, Hong, & Diamond, 1994) and Pictures tasks (Burns, Riggs, & Beck, 2012). The Day–Night task was converted into a computer task in order to measure response times (RTs) as well as accuracy.

Because the relationship between ToM and development of CR has not yet been studied, it is not clear which type of mental state representation might be involved. For this reason, a scale made up of three tasks was implemented in this study, based on Wellman and Liu's (2004) ToM scale, from which the following tasks were selected according to developmental suitability: knowledge access, contents false belief, and real apparent emotion. Age and receptive vocabulary were measured as control variables. ToM ability could be correlated with performance on CR tasks (i.e., ability to avoid interference from a partner and thereby perform well) in one of two ways: (a) a negative correlation because ToM allows children to represent a partner during a joint task, (b) a positive correlation because ToM helps children to separate representations of self and other and thus avoid interference.

To summarize, we investigated whether 4- to 5-year-olds' performance on CR tasks is correlated positively with IC and WM and correlated either positively or negatively with ToM.

Method

Participants

Participants were one hundred and fifteen 4- to 5-year-olds (M_{age} = 61.68 months, range = 48–69 months, 60 male) sampled from Birmingham's Think Tank Science Museum (n = 30), two from the primary schools in the Birmingham area (n = 18), and three from the Northamptonshire area (n = 67) between April 2013 and March 2014. The sample was from several schools with diverse intakes, consisting of approximately 90% Caucasian, 7% Asian/British Asian, and 3% Black/African/Carribbean children from working–middle class backgrounds (estimated from census data for each county, Office for National Statistics, 2011). Two participants from Northamptonshire schools failed to complete all tasks and were excluded from analyses. All other participants carried out all tasks and all conditions of each task in a within-subjects design.

Design and Procedure

Children were tested in a separate room, quiet corridor area, or reading corner of the schools or museum. Tasks were divided into two 15- to 20-min sessions in order to give children a rest between blocks. Depending on school classroom routines, for children in schools there was a period of a minimum of 1 hr to a maximum of 20 hr between completing the first block and starting the second session of tasks. Tasks consisted of one CR measure, three ToM measures (ToM), two WM

measures, and two IC measures, and the British Picture Vocabulary Scale (BPVS). The BPVS was included to make sure that any of the relationships of interest between CR, ToM, WM, and IC cannot be simply attributed to individual differences in language development. Tasks were administered in a fixed order. Session 1: bear/duck (CR), knowledge access (ToM), contents false belief (ToM), real apparent emotion (ToM), backward digit span (WM). Session 2: day–night Stroop (IC), counting and labeling (WM), pictures (IC); BPVS.

Measures

Receptive Vocabulary Measure

The BPVS, 2nd ed. (Dunn, Dunn, Whetton, & Burley, 1997) was fully administered to all children. This involved verbally presenting a word along with a card displaying four pictures, from which children were required to select the matching picture. Children were tested on a series of trials until reaching a failure criterion.

Working Memory Measures

Backward digit span. The methodology for this task was taken from Carlson et al. (2002). Children were introduced to a puppet called Ernie and told, "Ernie is being very silly. Everything I say, he says backwards! Like this, if I say the numbers '1, 2' he says '2, 1'! Do you think you can do what Ernie has done and say things backwards?" Participants then had a two-digit practice trial, on which they were corrected if they got the answer wrong, followed by three test trials with two, three, and four digits, respectively. Participants obtained a score out of four.

Counting and labeling. The methodology for this task was taken from Carlson et al. (2002), and the script is provided in Appendix S1. Children were introduced to a set of toys that they first had to count, then label, and finally count and label simultaneously.

Inhibitory control measures

Day–night Stroop. This task was adapted from the classic day–night Stroop task (Gerstadt et al., 1994) in a computerized version using EPrime 2 (Psychology Software Tools, Inc., Sharpsburg, PA, USA). Children were required to press a button with a picture of the moon and stars on when they heard a recorded voice say the word "day" and to

press a button with a picture of the sun on when they heard "night." Computer presentation made it possible to record both accuracy and RT. Participants completed 4 practice trials and 16 test trials. Trials commenced upon completion of the previous trial, giving no time limit on responses.

Pictures task. This task was identical to that used in Burns et al. (2012). Children participated in a computer game in which they saw a picture of either a monkey or a cat appear on the screen. They were required to press one of two buttons, one of which corresponded with the cat (and displayed a sticker with the same cat picture) and one with the monkey (displaying a monkey sticker). These buttons were placed 20 cm apart, with one on the left- and one on the right-hand side of the laptop screen. Pictures on the computer were also displayed either on the left- or right-hand side, so that half of the trials ($n = 10$) were congruent with the side on which the corresponding animal button was situated and half were incongruent. Participants completed a total of 4 practice and 20 test trials.

Theory-of-Mind Measures

ToM measures were taken from Wellman and Liu's (2004) ToM scale. Only three of the original tasks from the five-task scale were employed, due to the smaller age range being tested here and the age appropriateness of selected tasks.

Knowledge access. This task tests for understanding that others will not have knowledge of an event they were not present for. Participants were shown a Lego drawer and asked, "What do you think is inside the drawer?" The experimenter then opened the drawer to reveal a plastic monkey, "Let's see . . . Look, there's really a monkey inside!" She then closed the drawer, "Ok, so what is in the drawer?" After the child's response, the experimenter said, "Polly has never seen inside this drawer. Now here comes Polly." A plastic doll was then brought into view. The test question, "So, does Polly know what is in the drawer?" was then asked, followed by the memory check question, "Did Polly see inside the drawer?" In order to score correctly, children had to respond "no" to both of these questions.

Contents false belief. This tests understanding that others can have a belief that is different from reality. Children were presented with a plaster ("Band-Aid") box. They were first asked, "What do you think is in the plaster box?" after which the box was opened to reveal a toy pig, "Look, it's really a pig inside!" The box was then closed and the child was asked, "Ok, so what is in the plaster box?"

Once the child had answered, the experimenter said, "Peter has never seen inside this plaster box. Now here comes Peter." A different plastic doll was then brought into view and the test question was asked, "So, what does Peter think is in the box, plasters or a pig?" This was followed by the memory check question, "Did Peter see inside the box?" Children had to respond "plasters" and "no," respectively, in order to score correctly on this task.

Real apparent emotion. This tests understanding that individuals can portray different emotions from those which they are actually experiencing. Children were told a story about a boy who wanted to hide how he felt about an event. In the story, another child had called the boy names, which had made him feel sad. However, he tried to hide how he felt so that the other children would not think he was a baby (see Appendix S1 for full script). Children were asked a memory check question, "How did the boy really feel when everyone laughed, did he feel happy, sad, or ok?" They were given a sheet with three faces on (one happy, one neutral and one sad) which they could point to in response. They were then asked the test question, "How did he try to look on his face when everyone laughed, did he try to look happy, sad, or ok?" Children had to respond with a more negative emotion for the memory check question than for the test question in order to gain a score of 1.

Corepresentation task

This task was modified from the computer task employed in Milward et al. (2014, Experiment 3) in order to make it suitable for use within subjects. The task was a go–no-go task, adapted from the classic bear/dragon test of IC (Kochanska, Murray, Jacques, Koenig, & Vandegeest, 1996). Pairs of participants (always a child and the experimenter) each had a response button that they had to press in response to one of two stimuli according to a given rule. For example, one might have to respond to a picture of a duck but inhibit responding to a bear. In contrast to Milward et al. (2014), the current task consisted of two rather than three blocks of 12 bear and duck trials. This was in order to allow an equal number of blocks before and after the switch. A second EPrime program was designed that had an identical structure to the first program, but rather than presenting bear and duck stimuli, they presented pictures of a pig and a penguin (see Figure S1). This was so that representations of stimuli from the first (same)

task could not influence those in the second (different) task, and to aid children's understanding of the different rules for each task.

Pilot data found an effect of task type (same vs. different) in a within-subjects design only when the same condition was presented first and the different condition was presented second. This is unlikely to have been caused solely by children getting tired or bored and thus declining in performance over time, because if this were the case, then such a decline should also occur when the same condition was presented second. One possibility was that there would be a carryover effect when the different task was presented first, in that children continued to represent the other task in the same task condition having done so in the preceding different task condition. For this reason, tasks were presented in a fixed order with the same condition first and then different condition second.

Each child completed both conditions with a different version of the EPrime program (bear/duck or pig/penguin) for each condition. The order of presentation of each program was counterbalanced so that half of the participants completed the bear/duck version in the same condition and half in the different condition, and likewise for the pig/penguin version. Children completed the solo practice, joint practice, and same condition as in Milward et al. (2014). Once the same condition had been completed, children were asked a comprehension question and then given the instructions for the different condition (example given is for same task with bear as target, different task with pig as child's target), "Now we're going to do something different. This time your job is to spot the Pig. So whenever you see the Pig, you should press your button as fast as you can. But, if you see the Penguin, you shouldn't press your button. You just stay still and don't press anything at all. My job is to spot the Penguin. So whenever I see the Penguin I'm going to press my button as fast as I can. But, if I see the Pig, I'm not going to press my button, I'm just going to stay still and not press anything at all." A further joint practice was then administered using the new stimulus set and new rules. The second EPrime program was then loaded and children completed the different condition, followed by a comprehension question.

Results and Discussion

Descriptive statistics for all tasks are presented in Table 1. Results will be presented in two stages.

Table 1
Descriptive Statistics for All Tasks

Task	Mean score (SD)	
	Accuracy	RT
Receptive vocabulary		
BPVS raw score	54.12 (11.24)	
Working memory		
Backward digit span (0–4)	1.63 (1.40)	
Counting and labeling (0–2)	0.95 (0.94)	
Total correct answers (0–6)	2.59 (1.99)	
Inhibitory control		
Pictures: congruent	94% (11)	1,240.94 (204.83)
Pictures: incongruent	90% (13)	1,393.87 (210.70)
Day–night Stroop	86% (25)	1,618.49 (325.50)
Theory of mind		
Knowledge access	82% (38)	
Contents false belief	57% (50)	
Real apparent emotion	27% (45)	
Total correct answers (0–3)	1.66 (0.90)	
Corepresentation		
Same task	85% (11)	727.32 (116.90)
Different task	83% (13)	751.25 (101.30)

Note. RTs were analyzed only for correct responses. BPVS = British Picture Vocabulary Scale; RTs = response times.

First, composite measures will be described and correlational analyses carried out to identify relations between all measures. Regression analyses will then be carried out in order to identify the contribution of factors to CR.

Composite Measures

Inhibitory Control

Day–night Stroop. Accuracy and RTs were recorded for this task. RTs were analyzed only for correct responses. Responses with RTs below 250 ms and above 3 *SD* from the mean (6,335.77 ms) were excluded. In order to combine accuracy and RTs in a composite measure for performance, values were converted into *z* scores and signs were reversed for RT *z* scores so that, as for accuracy, higher scores signified better performance. *z* Scores for accuracy and RTs were then added together. This was as an alternative to inverse efficiency scores, which also combine information from accuracy and RT, but which are argued to be unsuitable for scores with error rates over 10% (Bruyer & Brysbaert, 2011). Because the task always required IC, the composite performance measure is the inhibition ability score. The higher

the performance, the stronger the inhibition ability. The composite performance score for this task will be referred to as "day–night composite" henceforth.

Pictures. Accuracy and RTs were recorded for both congruent and incongruent conditions of this task. RTs were analyzed only for correct responses. A measure of the inhibitory ability of this task was calculated in three steps. First, for each condition, we calculated the composite performance measure, based on both accuracy and RT, in the same way as for the day–night Stroop. In this performance measure, a higher value indicates better performance. Second, the composite performance measure in the congruent condition was regressed from that in the incongruent condition. Third, the inhibition ability scores were calculated as residuals for this regression model. That is, the inhibition ability of a given participant was measured as how much his or her performance in the incongruent condition deviated from that expected from his or her performance in the congruent condition, based on the regression model; thus, a bigger positive residual indicates a better inhibition ability. This inhibition ability score will be referred to as "picture residual score" henceforth.

This approach was first proposed by DeGutis, Wilmer, Mercado, and Cohan (2013). The method uses residuals to calculate the variance of a factor of interest (inhibition ability) that is present in the experimental condition (the incongruent case) but not in the control condition (the congruent condition) while controlling for variance shared between the two conditions. The advantage of this method can be illustrated by the fact that residuals (inhibition ability) do not correlate with the performance in the control condition (the congruent condition). A larger value for picture residual score indicates a better inhibition ability. The residual method was used as an alternative to the subtraction method, where the performance in the control condition (congruent condition) is subtracted from that in the experimental condition (incongruent condition) in order to produce a measure of a factor of interest (inhibition ability). The weakness of the subtraction method can be illustrated by the fact that the resulting difference score correlates with the performance in the control condition. That is, the difference score is "contaminated" by factors relevant for the performance of the control condition (congruent condition).

Corepresentation. Accuracy and RTs were recorded for same and different conditions. There was a descriptive difference between the two conditions for both accuracy and RTs (see Table 1).

Paired samples *t* tests highlighted a significant difference between conditions for correct RTs, *t*(107) = −2.56, *p* = .01, but not for accuracy, *t*(111) = 1.66, *p* = .10. This replicates the effect of task type in RTs in a previous version of this task (Experiment 3 in Milward et al., 2014) that found a between-subjects difference in task type, and pilot data that found a within-subjects difference. In order to create a composite score for CR, RT, and accuracy values for each condition were converted into *z* scores and added together within each condition to create a speed–accuracy composite (reversing signs for RTs so that higher values indicate better performance). This gave a mean of .86 (*SD* = 1.06) for the same task and .84 (*SD* = 1.08) for the different task. Next, residuals were calculated in the same way as for the pictures task (following DeGutis et al., 2013), by regressing the same task composite from the different task composite and calculating the inhibition ability score as residuals for this regression model. Variance that is unique to the different condition is the most interesting, as only in this condition CR is expected to show an effect of conflict with a partner. The residual score will be referred to as "CR residual score" henceforth (the higher the score, the better the performance and lower the interference from CR). For readers interested in descriptive statistics and correlations for composite scores prior to residual calculations (see Table S1). These may be more indicative of general task requirements in this type of joint task rather than those specific to CR interference.

Correlations. First, where there were multiple tasks within a single domain (WM, IC, and ToM), task means were entered into Pearson's correlational analyses. Note that a higher value indicates better performance in BPVS, ToM scores, WM scores, and the inhibition score (better inhibition ability). Backward digit span and counting and labeling correlated significantly, *r*(113) = .43, *p* < .001, and were therefore added together to create an overall WM score ("working memory total," henceforth). Day–night Stroop and pictures residual score did not correlate, *r*(111) = .02, *p* = .81. For this reason, further analyses included both a total inhibition score (by adding day–night Stroop composite and pictures residual score, "inhibition total" henceforth) as well as each of the individual inhibition measures independently. Of the three ToM tasks, only knowledge access and contents false belief correlated significantly, *r*(111) = .39, *p* < .001. For this reason, the three ToM tasks were analyzed separately as well as including a total ToM score by adding all three task scores ("ToM total," henceforth).

Age and receptive vocabulary correlations. An overall Pearson's correlation matrix can be seen in Table 2. Age in months correlated with receptive vocabulary, WM total, and day–night composite. Receptive vocabulary correlated with all three ToM measures, ToM total, WM total, day–night composite, and CR. This is consistent with existing literature demonstrating the development of WM, inhibition, and ToM at this age (Carlson, 2010; Carlson et al., 2002).

Executive function and ToM correlations. WM total and inhibition total correlated with each other significantly, *r*(107) = .26, *p* = .006. This remained significant after age and receptive vocabulary had been partialed out, *r*(104) = .20, *p* = .04. This correlation was descriptively weaker than that found between the two WM measures, *r*(113) = .43, *p* < .001, which is consistent with Miyake et al.'s (2000) prediction that executive function measures should correlate with one another but more strongly between measures of the same type of executive component. However, inhibition measures did not correlate with one another. This was unexpected, given that they are both conflict inhibition tasks, where participants are required to inhibit a conflicting stimulus in order to respond to a target.

One possible explanation is that the day–night task in this study did not feature congruent trials as a control measure, whereas the pictures task did so. Therefore, the composite scores for each task are slightly different, in that the day–night composite does not control for factors that may be involved in performance beyond IC. Additionally, whereas the day–night composite is determined by a participant's ability to retain in mind and act upon a rule that conflicts with a prepotent response, the pictures task measures ability to switch between inhibiting a prepotent response on incongruent trials while acting consistently with it on congruent trials. Furthermore, the competing element in the pictures task is spatial, whereas the day–night is conceptual. In conclusion, there are reasons to explain the lack of correlation here, although it was unexpected given previous findings that within domain executive function tasks are correlated (Miyake et al., 2000).

Two of the ToM measures correlated with WM total, knowledge access: *r*(111) = .34, *p* < .001; contents false belief: *r*(111) = .35, *p* < .001. These results remained similar once age in months and receptive vocabulary had been partialed out (significant

Table 2

R Values for Pearson's Correlations Between Age, Receptive Vocabulary, ToM, Working Memory, Inhibition, and Corepresentation

	Age (months)	BPVS raw score	Knowledge access	Contents false belief	Real apparent emotion	ToM total	Working memory total	Day–night composite	Pictures residual	Inhibition total	Corepresentation residual
Age (months)	—	.29**	.10	.09	.05	.11	.35**	.22*	-.04	.16	.05
BPVS raw score		—	.29**	.34**	.21*	.41**	.48**	.25**	-.08	.15	.17
ToM											
Knowledge access			—	.39**	.03	.65**	.34**	.04	-.05	.00	.19*
Contents false belief				—	.14	.79**	.35**	.17	-.05	-.11	.17
Real apparent emotion					—	.58**	.12	.07	.14	.13	.16
Total						—	.40**	.15	.02	.13	.26**
Working memory											
Total							—	.36**	-.08	.26**	.20*
Inhibition											
Day–night composite								—	.02	.88**	.36**
Pictures residual									—	.49**	-.08
Total										—	.27**
Corepresentation											
Corepresentation residual											—

Note. Higher values indicates better performance on the task. RTs were analyzed only for correct responses. BPVS = British Picture Vocabulary Scale; ToM = theory of mind; RTs = response times. *$p < .05$. **$p < .01$.

correlations between WM total and knowledge access, $r(104) = .25$, $p = .01$, and between pictures residual score and real apparent emotion, $r(104) = .21$, $p = .03$.

CR and cognitive measures. The CR residual score correlated with knowledge access, $r(106) = .19$, $p < .05$, ToM total, $r(106) = .26$, $p < .01$, WM total, $r(106) = .20$, $p < .04$, day–night composite, $r(106) = .36$, $p = .001$, and inhibition total, $r(101) = .27$, $p < .01$. Regression analyses were carried out in order to identify which relationships remained significant predictors once other variables were controlled for.

Regression analyses. The CR residual score was used as the dependent variable. This score measures the unique variance in the performance of the different condition, after controlling for the performance in the same condition (DeGutis et al., 2013). Only in the different condition, CR is expected to show an effect of conflict with a partner's task; thus, this score is a good measure of the strength of CR (the lower the score, the stronger the interference effect due to CR of the partner).

A multiple linear regression was carried out with CR residual score as the dependent variable and age in months, BPVS raw score, inhibition total, WM total, and ToM total as the predictor variables in a single model using enter method (see Table 3). The overall model was significant, $F(5, 97) = 3.63$, $p = .005$, with inhibition total and ToM total as the only significant predictors within the model (inhibition total: $\beta = .12$, $t = 2.34$, $p = .02$; ToM total: $\beta = .20$, $t = 2.13$, $p = .04$).

Two hierarchical regressions were subsequently carried out in order to obtain R^2 values for each of these significant predictors. The first regression was identical to the initial regression model but entering

Table 3
Regression Output for Multiple Regression With Corepresentation Residual Score as the Dependent Variable and Age in Months, BPVS Raw Score, ToM Total, Working Memory Total, and Inhibitory Control Total as Predictors

	B	SE	β	t	Sig.
(Constant)	−.305	1.031		−0.296	.768
Age in months	−.007	0.017	−.044	−0.432	.666
BPVS raw	.006	0.009	.077	0.690	.492
ToM total	.207	0.097	.223	2.133	.035
Working memory total	.030	0.048	.071	0.623	.535
Inhibition total	.122	0.052	.228	2.339	.021

Note. BPVS = British Picture Vocabulary Scale; ToM = theory of mind.

inhibition total in a second step (R^2 change = .047, $p = .02$). The second was identical to the initial model but entering ToM total in a second step (R^2 change = .04, $p = .035$). The positive significant βs and R^2 values indicate that both Inhibition and ToM contribute independently to individual scores on the CR residual score. More specifically, participants who have stronger inhibition abilities and stronger ToM abilities showed better performance on (i.e., less interference from the joint action partner's role in) the CR task.

General Discussion

We found independent roles of both ToM and IC in performance on the CR task as measured by residuals between same and different tasks. More specifically, ToM and IC each had an independent positive relationship with the ability to avoid interference from a partner in a joint task (i.e., less conflict in the different task condition of the CR task). This suggests that ToM and IC are involved in joint action in 4- to 5-year-olds. Additionally, although WM correlated with the CR task, this was not found to contribute once other factors had been included in a regression model. This, along with the fact that age, receptive vocabulary, and WM were controlled for in these analyses, suggests a specific role of ToM and IC, which goes beyond a more general maturation in cognitive ability that would be shared between all of these measures. These findings will be discussed in the following paragraphs.

The finding regarding IC is consistent with the prediction that this ability is required in order to avoid interference caused by corepresenting a partner in the different task condition. Participants with stronger inhibition abilities performed better on the different task once controlling for the same task. The fact that this effect was independent of ToM score suggests that general executive abilities are additionally involved in corepresenting a partner during a joint task, beyond any specifically social abilities measured by the ToM task. It is important to distinguish such general inhibitory processes from self–other conflict resolution processes that are specific to ToM and have been shown to depend on different brain regions to those used for general IC (Samson, Houthuys, & Humphreys, 2015).

The finding regarding ToM is consistent with the prediction that ToM helps children to perform better on CR tasks because it helps them to clearly differentiate and separate task representations of self

and other. This makes sense if one considers the nature of the tasks implemented in the ToM scale used here. All three of these tasks involve being able to separate one's own mental state from that of a protagonist. In the knowledge access task, one's own state of knowing must be separated from that of a protagonists' state of naivety. This is similar for the contents false belief task, although here the content of one's own belief must be separated from the incorrect belief of the other. In the real apparent emotion task, one needs to separate one's own knowledge of the protagonist's true emotion from that portrayed on his face. Many frameworks for ToM have suggested that it is not a unitary construct but rather involves separate processes for construction of representations of self and other versus resolving conflict between these representations (e.g., Apperly, 2010; Carlson & Moses, 2001; Doherty, 2008; Frith & Frith, 2006; Leslie & Thaiss, 1992). The current results provide evidence that the latter ToM process is involved in CR, enabling conflict resolution between one's own and one's partner's task. Note that this analysis applies even when the "other" perspective is actually one's past self, as in the representational change task (Astington & Gopnik, 1988) in which children find it as difficult to report their own prior false belief as they do someone else's current false belief. However, it is not clear whether the same analysis applies to all tasks involving differences between self and other, which vary both in the age at which children first succeed, and in the nature of the perspective problem that is posed (see Rakoczy, 2010 for a discussion).

It is possible that this finding of a role of self–other distinction is due to it being a shared mechanism between multiple social domains, including ToM and CR. This is an interesting avenue to investigate further in the context of joint action, which may require processing of multiple types of social information from motor, cognitive, and even affective domains in unison, and therefore could benefit from a shared self–other distinction mechanism across these domains. It should be noted, however, that the conclusion that the self–other distinction component of ToM is involved in CR does not rule out the possibility that the representation construction component may not play an additional role, perhaps at a younger age than tested here. This should be addressed in future research.

One possibility for future research is to investigate the role of set shifting on CR. Although considered less theoretically relevant than conflict IC in the current study, there is perhaps reason to consider it in the future. This is because some researchers have suggested that CR may be a mechanism for turn taking in joint action (Wenke et al., 2011), whereby participants represent not *what* their partner is doing but *whose* turn it is to act. In this case, one might not need to activate one rule about a stimulus–response sequence ("bear" = go) and inhibit the other ("duck" = no-go), but rather switch between two rules about whose turn it is to act ("bear" = my turn" vs. "duck" = your turn). If set shifting did contribute independently beyond conflict IC, it might shed some light on this argument.

The current findings suggest that higher level abilities such as ToM and IC are involved in joint action by 4- to 5-year-olds. This finding is important in light of the discussion of how to interpret children's ability to engage in joint action (Brownell, 2011). Some (e.g., Warneken et al., 2012) argue that children's joint action is underpinned by higher level cognitive abilities to represent partners, but others (e.g., Butterfill, 2011 and our discussion in the introduction of the current article) cautioned that accounts based on low-level cognitive abilities may be possible for previous findings. The study reported here suggests that at least at this age, children do employ higher level abilities. Clearly it is possible that lower level skills can provide children with the means to participate in activities that may be considered joint in some important respects at a younger age. For example, even 12-month-olds will choose social partners based on their degree of movement synchrony with the infant (Tunçgenç, Cohen, & Fawcett, 2015), suggesting a sensitivity to interpersonal relations that could result in active selection of desirable action partners. Goal tracking may also be an earlier-developing skill that helps children to engage jointly with others (Butterfill, 2011). Preschoolers may also understand the potential of a partner to be used as a social tool to achieve their own goal (Warneken et al., 2006, 2012). However, these mechanisms do not necessarily require any representational capacity for self and other in unison and therefore would not fulfill a conservative definition of joint action, such as in Bratman (1992). If claims for high-level explanations are to be substantiated, future research needs to focus on finding evidence for involvement of such representational capacities in joint action. Individual differences studies looking at the relationship between different types of joint action, and ToM measures could be one way of doing this.

The current findings leave open the possibility that the representation construction component of ToM plays an additional role in emergence of CR

itself, which would explain the age of onset of CR effects (4- to 5-year-olds, Milward et al., 2014) and would be consistent with previous studies suggesting that ToM ability might be necessary for CR interference to occur (Humphreys & Bedford, 2011). Thus, further research should distinguish between three possibilities as to the emergence of CR: (a) that the representation construction component of ToM is a necessary condition for CR of a partner to emerge, but that subsequent development of the self–other distinction component helps to avoid interfering effects on performance from representation of the other; (b) that alternative cognitive abilities that develop at around 4–5 years of age can explain the lower bound on CR in development (4- to 5-year-olds, but not in 2- to 3-year-olds; Milward et al., 2014); or (c) CR develops earlier than has been found in existing paradigms, and could be measured at a younger age with more sensitive tasks. Such future research might open up further opportunities for investigating the role in CR of other developing cognitive abilities that have been considered relevant to general joint action development, such as self–other understanding (Brownell & Carriger, 1990), intentionality and joint goal understanding (Butterfill, 2011; Warneken et al., 2012), and lower level perception–action links (Di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Meltzoff & Moore, 1977), all of which could be studied using an individual differences methodology such as has been employed here. By identifying which of these established processes or mechanisms are related to performance on the various joint action paradigms that have been studied to date, researchers can build a better picture of what underlies these behaviors and consequently how we should define them.

To conclude, the current results showed that individual differences in the strength of CR effect in 4- to 5-year-olds was predicted independently by children's ToM ability and by their IC ability. Those with better abilities in ToM and in IC had less interference from the joint action partner's role in the CR task. We argued that ToM aids children in forming separate, stable representations of self and other, whereas inhibition helps avoid the resulting interference caused by representation of multiple task rules.

Acknowledgments

References

Apperly, I. A. (2010). *Mindreaders: The cognitive basis of a "theory of mind"*. Hove, UK and New York, NY: Psychology Press.

Astington, J. W., & Gopnik, A. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, *59*, 26–37. doi:10.2307/1130386

Atmaca, S., Sebanz, N., & Knoblich, G. (2011). The joint flanker effect: Sharing tasks with real and imagined co-actors. *Experimental Brain Research*, *211*, 371–385. doi:10.1007/s00221-011-2709-9

Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling versus unable: Infants' understanding of intentional action. *Developmental Psychology*, *41*, 328–337. doi:10.1037/0012-1649.41.2.328

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*, 647–663. doi:10.1111/j.1467-8624.2007.01019.x

Bratman, M. E. (1992). Shared cooperative activity. *The Philosophical Review*, *101*, 327–341. doi:10.2307/2185537

Brownell, C. A. (2011). Early developments in joint action. *Review of Philosophy and Psychology*, *2*, 193–211. doi:10.1007/s13164-011-0056-1

Brownell, C., & Carriger, M. (1990). Changes in cooperation and self-other differentiation during the second year. *Child Development*, *61*, 1164–1174. doi:10.2307/1130884

Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychologica Belgica*, *51*, 5–13. doi:10.5334/pb-51-1-5

Burns, P., Riggs, K. J., & Beck, S. R. (2012). Executive control and the experience of regret. *Journal of Experimental Child Psychology*, *111*, 501–515. doi:10.1016/j.jecp.2011.10.003

Butterfill, S. A. (2011). Joint action and development. *Philosophical Quarterly*, *62*, 23–47. doi:10.1111/j.1467-9213.2011.00005.x

Carlson, S. M. (2010). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, *28*, 595–616. doi:10.1207/s15326942dn2802_3

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*, 1032–1053. doi:10.1111/1467-8624.00333

Carlson, S. 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*, 73–92. doi:10.1002/icd.298

Carpenter, M. (2009). Just how joint is joint action in infancy? *Topics in Cognitive Science*, *1*, 380–392. doi:10.1111/j.1756-8765.2009.01026.x

Decety, J., & Lamm, C. (2007). The role of the right termporoparietal junction in social interaction: How low-level computational processes contribute to meta-cognition. *Neuroscientist*, *13*, 580–593. doi:10.1177/1073858407304654

DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition*, *126*, 87–100. doi:10.1016/j.cognition.2012.09.004

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91*, 176–180. doi:10.1007/BF00230027

Doherty, M. J. (2008). *Theory of mind: How children understand others' thoughts and feelings*. Hove, UK: Psychology Press.

Dolk, T., Hommel, B., Prinz, W., & Liepelt, R. (2013). The (not so) social Simon effect: A referential coding account. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 1248–1260. doi:10.1037/a0031031

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale* (2nd ed.). Windsor, UK: NFER-Nelson.

Frith, C. D., & Frith, U. (2006). The neural basis of mentalising. *Neuron*, *50*, 531–534. doi:10.1016/j.neuron.2006.05.001

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3.5–7 years old on a Stroop-like day-night test. *Cognition*, *53*, 129–153. doi:10.1016/0010-0277(94)90068-X

Hay, D. F. (1979). Cooperative interactions and sharing between very young children and their parents. *Developmental Psychology*, *15*, 647–653. doi:10.1037/0012-1649.15.6.647

Humphreys, G. W., & Bedford, J. (2011). The relations between joint action and theory of mind: A neuropsychological analysis. *Experimental Brain Research*, *211*, 357–369. doi:10.1007/s00221-011-2643-x

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development*, *67*, 490–507. doi:10.2307/1131828

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition*, *43*, 225–251. doi:10.1016/0010-0277(92)90013-8

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). *Developmental Science*, *7*(3), 297–307.

Liebal, K., Behne, T., Carpenter, M., & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science*, *12*, 264–271. doi:10.1111/j.1467-7687.2008.00758.x

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 75–78. doi:10.1126/science.198.4312.75

Milward, S., Kita, S., & Apperly, I. A. (2014). The development of co-representation effects in a joint task: Do children represent a co-actor? *Cognition*, *132*, 269–279. doi:10.1016/j.cognition.2014.04.008

Miyake, A., Friedman, N. P., Emerson, M. J., Witzi, A. H., & Howerter, A. (2000). The unity and diversity of Executive Functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100. doi:10.1006/cogp.1999.0734

Moll, H., Richter, N., Carpenter, M., & Tomasello, M. (2008). Fourteen-month-olds know what "we" have shared in a special way. *Infancy*, *13*, 90–101. doi:10.1080/15250000701779402

Office for National Statistics. (2011). *Census data*. Retrieved from http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/index.html

Perner, J. (1991). *Understanding the representational mind. Learning, development, and conceptual change*. Cambridge, MA, US.

Rakoczy, H. (2010). Executive function and the development of belief-desire psychology. *Developmental Science*, *13*, 648–661. doi:10.1111/j.1467-7687.2009.00922.x

Saby, J. N., Bouquet, C. A., & Marshall, P. J. (2014). Young children co-represent a partner's task: Evidence for a joint Simon effect in five-year-olds. *Cognitive Development*, *32*, 38–45. doi:10.1016/j.cogdev.2014.08.001

Samson, D., Houthuys, S., & Humphreys, G. W. (2015). Self-perspective inhibition deficits cannot be explained by general executive control difficulties. *Cortex*, *70*, 189–201. doi:10.1016/j.cortex.2014.12.021

Santiesteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., & Bird, G. (2012). Training social cognition: From imitation to theory of mind. *Cognition*, *122*, 228–235. doi:10.1016/j.cognition.2011.11.004

Schuwerk, T., Döhnel, K., Sodian, B., Keck, I. R., Rupprecht, R., & Sommer, M. (2014). Functional activity and effective connectivity of the posterior medial prefrontal cortex during processing of incongruent mental states. *Human Brain Mapping*, *35*, 2950–2965. doi:10.1002/hbm.22377

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others: Just like one's own? *Cognition*, *88*, 11–21. doi:10.1016/S0010-0277(03)00043-X

Sowden, S., & Catmur, C. (2015). The role of the right temporoparietal junction in the control of imitation. *Cerebral Cortex*, *25*, 1107–1113. doi:10.1093/cercor/bht306

Spengler, S., von Cramon, D. Y., & Brass, M. (2009). Control of Shared Representations Relies on Key Processes Involved in Mental State Attribution. *Human Brain Mapping*, *30*, 3704–3718.

Tsai, C., Kuo, W., Hung, D. L., & Tzeng, O. J. L. (2008). Action co-representation is tuned to other humans. *Journal of Cognitive Neuroscience*, *20*, 2015–2024. doi:10.1162/jocn.2008.20144

Tunçgenç, B., Cohen, E., & Fawcett, C. (2015). Rock with me: The role of movement synchrony in infants' social

and nonsocial choices. *Child Development*, *86*, 976–984. doi:10.1111/cdev.12354

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, *77*, 640–663. doi:10.1111/j.1467-8624.2006.00895.x

Warneken, F., Gräfenhain, M., & Tomasello, M. (2012). Collaborative partner or social tool? New evidence for young children's understanding of joint intentions in collaborative activities. *Developmental Science*, *15*, 54–61. doi:10.1111/j.1467-7687.2011.01107.x

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *3*, 1301–1303. doi:10.1126/science.1121448

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*, 523–541. http://www.jstor.org/stable/3696656

Wenke, D., Atmaca, S., Höllander, A., Liepelt, R., Baess, P., & Prinz, W. (2011). What is shared in joint action? Issues of co-representation, response conflict, and agent identification. *Review of Philosophy and Psychology*, *2*, 147–172. doi:10.1007/s13164-011-0057-0

Wimmer, H. H., & Perner, J. J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children. *Cognition*, *13*, 103. doi:10.1016/0010-0277(83)90004-5

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Figure S1. Pig and Penguin Stimuli for Corepresentation Task

Table S1. Correlation Matrix for Composite Speed–Accuracy Measures of Corepresentation (Same and Different Tasks) With Theory of Mind, Inhibitory Control, and Working Memory

Appendix S1. Scripts for Counting and Labeling and Real Apparent Emotion