



Three-year-olds' Perspective-taking in Social Interactions: Relations with Socio-cognitive Skills

Natalie Brezack, Marlene Meyer & Amanda L. Woodward

To cite this article: Natalie Brezack, Marlene Meyer & Amanda L. Woodward (2021): Three-year-olds' Perspective-taking in Social Interactions: Relations with Socio-cognitive Skills, Journal of Cognition and Development, DOI: [10.1080/15248372.2021.1901713](https://doi.org/10.1080/15248372.2021.1901713)

To link to this article: <https://doi.org/10.1080/15248372.2021.1901713>



© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.



[View supplementary material](#)



Published online: 15 Apr 2021.



[Submit your article to this journal](#)



[View related articles](#)



[View Crossmark data](#)

Three-year-olds' Perspective-taking in Social Interactions: Relations with Socio-cognitive Skills

Natalie Brezack ^a, Marlene Meyer ^{a,b*}, and Amanda L. Woodward ^a

^aUniversity of Chicago, USA; ^bRadboud University Nijmegen, The Netherlands

ABSTRACT


Understanding others' perspectives and integrating this knowledge in social interactions is challenging for young children; even adults struggle with this skill. While young children show the capacity to understand what others can and cannot see under supportive laboratory conditions, more research is necessary to understand how children implement their perspective-taking (PT) skill during interactions and which socio-cognitive skills support their ability to do so. This preregistered study examined children's Level 1 visual PT in a real-time social interaction and tested whether social-cognitive skills (focusing on inhibition of imitation) predicted PT. Thirty-six 3-year-old children (mean age: 37.3 months) participated in a PT task and responded implicitly (via eye gaze) and explicitly (via toy choice) to situations where their communicative partner could see some objects but not others. Three-year-olds demonstrated sensitivity to another's perspective via implicit responses, but did not consistently take their partner's perspective into account in their actions when considering objects their partner could not see. Contrary to adult findings, children who struggled to inhibit imitating (those more affected by another's actions) demonstrated better PT, again when considering objects outside their partner's sight. Thus, 3-year-olds' sensitivity to others' perspectives was robust, while *acting* on PT knowledge may still be developing; further, children more affected by another's actions demonstrated improved PT skills.

Introduction

Considering the perspective of another person is vital for successful communicative and social interactions. Imagine on her way out of the house, your friend says, "Toss me those keys!" You see two keyrings on the table and infer that she must be talking about the keys she can see, not the keys hidden from her view behind a vase of flowers. Such exchanges requiring analysis of another's perspective followed by timely action based on that analysis are staples of everyday social interactions. From a young age, children can consider another's perspective, but children and even adults sometimes struggle to account for others' points of view (Keysar, Barr, Balin, & Brauner, 2000). This study examined the question: How does three-year-olds' visual perspective-taking (PT) unfold in real-time social interactions and which social-cognitive skills support PT?

CONTACT Natalie Brezack  nbrezack@uchicago.edu  Department of Psychology, University of Chicago, 5848 South University Avenue, Chicago, IL 60637.

*Natalie Brezack and Marlene Meyer should be considered joint first author.

 Supplemental data for this article can be accessed on the [publisher's website](#).

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

The simplest and earliest-emerging form of PT is the understanding that what you see may differ from what someone else sees (“Level 1 Visual PT”; Masangkay et al., 1974; Moll & Tomasello, 2006). Early in life, infants and toddlers can engage in PT, analyzing another’s perspective. For example, infants follow another person’s gaze only when the gaze path is free of objects that could block where someone is looking (Caron, Kiel, Dayton, & Butler, 2002; Dunphy-Lelii & Wellman, 2004). Infants also attribute goal-directed reaching to people when target objects are visible to the agent (Luo & Baillargeon, 2007; Sodian, Thoermer, & Metz, 2007). This suggests that infants understand that others cannot see through solid materials, demonstrating early-emerging precursors to PT.

PT in social interactions

Beyond understanding another’s perspective, children need to use this knowledge to contribute successfully to social interactions. Analogous to quickly inferring your friend’s perspective on the keyring, children need to consider and act on others’ points of view to smoothly engage in communicative social interactions. Moll and Tomasello (2006) demonstrated that two-year-old children can use Level 1 PT knowledge in communicative contexts. In this study, toddlers saw an experimenter and two objects. Children could always see both objects, but one object was occluded from the experimenter’s view. In the experimental condition, the experimenter searched for an object and exclaimed that he could not find it. Children inferred that the object the experimenter was looking for must be the toy occluded from his view; 24-month-olds, but not 18-month-olds, reliably handed the experimenter the occluded object, successfully acting on their PT knowledge to provide the experimenter with the requested toy. Yet, their performance was far from perfect and seemed to depend on the social support available in the task. In Moll and Tomasello’s (2006) control condition, the experimenter simply requested a toy, and children at both ages handed the hidden object at chance levels. This suggests that additional prompting may be necessary to encourage children to think about others’ perspectives, at least early in life. While this study provides evidence that toddlers *can* engage in PT, the findings also indicate that children do not necessarily do so robustly.

Indeed, it can be a challenge for children and even adults to engage in PT during communicative interactions. In tasks similar to those of Moll and Tomasello (2006), toddlers demonstrated chance-level performance when considering a referent object their social partner either could or could not see (Herold & Akhtar, 2008; Krogh-Jespersen, Liberman, & Woodward, 2015; Liberman, Woodward, Keysar, & Kinzler, 2017). Three-year-old children also struggled to consider when another person could or could not see them, such as when playing hide-and-seek (Peskin & Ardino, 2003). Even adults sometimes fail to fully account for another’s point of view. For example, in the Director’s Task paradigm, adult participants moved objects in a grid array following instructions of a social partner (or “director”) who shared the participant’s view of some but not all objects in the array. Sometimes, the director’s instructions could refer to two objects, one of which was occluded from the director’s view. In these referentially ambiguous cases, participants needed to account for the director’s lack of view on one object to respond correctly by instead selecting the jointly perceived object. However, adults often attempted to move the object the director could not see, suggesting that even adults make errors when analyzing another’s perspective (Keysar et al., 2000). As children often fail PT tasks and adults make mistakes,

both children and adults have been described as “egocentric” in their communication (Epley, Morewedge, & Keysar, 2004). Despite these claims, other work has demonstrated that children can consider others’ points of view in non-egocentric ways (e.g., Khu, Chambers, & Graham, 2020). Clearly, more research is necessary to examine young children’s PT understanding in social interactions.

Prior studies with young children have largely focused on finding evidence that children can succeed at PT, rather than considering the conditions under which it is easier or harder for them to do so. In the current study, we focused on the latter issue. We evaluated 3-year-old children’s PT as revealed in their implicit responses in shifting attention to requested objects as well as their explicit responses to a communicative partner in a PT task. Studies have shown that infants and toddlers are sensitive to others’ perspectives and knowledge states when measured implicitly via eye gaze, before their explicit responses (speech, pointing, action) are evident (Baillargeon, Scott, & He, 2010). Much of this past research has focused on young children’s understanding of others’ false beliefs (one element of Theory of Mind). While children cannot verbalize responses to false belief tasks until about age four (Baron-Cohen, Leslie, & Frith, 1985), their eye gaze (Onishi & Baillargeon, 2005; Scott, He, Baillargeon, & Cummins, 2012) and spontaneous behaviors (Buttelmann, Carpenter, & Tomasello, 2009) reveal false belief understanding in infancy and toddlerhood. More analogous research would benefit our understanding of how young children analyze others’ visual perspectives.

Such research has been conducted with older children by examining their integration of others’ perspectives in referential communication tasks. Four- to 6-year-olds accounted for another’s perspective both when measured explicitly as children made statements to guide a partner’s toy choice and implicitly as children looked when responding to a partner’s toy request (Nadig & Sedivy, 2002; Bahtiyar & Küntay, 2009; Nilsen & Graham, 2009: Experiment 1). Three-and-a-half- to 4-year-olds similarly demonstrated PT understanding via both explicit (toy choice) and implicit (eye gaze) responses to objects a social partner could see, though children’s analyses of others’ perspectives were initially more egocentric and improved with age (Nilsen & Graham, 2009: Experiment 2). Implicit responses to another person’s perspective are rapid, immediately reflected in children’s gaze (Khu et al., 2020; San Juan, Khu, & Graham, 2015). However, some studies show dissociations between children’s looking and behavioral response to others’ perspectives; children sometimes demonstrate implicit responses indicating sensitivity to another’s perspective without integrating that sensitivity into their actions (Nilsen et al., 2008). To better understand whether younger children similarly show PT sensitivity when measured implicitly and explicitly, we tested 3-year-olds’ PT skills implicitly via eye gaze as children analyzed another’s perspective and explicitly via behavioral toy choice as children acted on this analysis. In line with research on infants’ understanding of others’ false beliefs and older children’s referential communication abilities, we predicted that young children would demonstrate better PT skills in their implicit responses before PT skills were reflected in their explicit responses.

We also tested PT under conditions that required children to consider what a social partner could see as well as what she could not see. Although PT is required in both situations, there is evidence that determining what others *cannot see* may be particularly challenging. School-aged children and adults were faster to reason about objects that were jointly viewed compared to those that only the participant could see (McCleery, Surtees,

Graham, Richards, & Apperly, 2011; Surtees & Apperly, 2012). Yet, studies on the early development of PT have separately investigated whether infants and toddlers understand what someone can see (Krogh-Jespersen et al., 2015) or cannot see (Moll & Tomasello, 2006). To our knowledge, no study has evaluated both types of PT judgments within a single paradigm or investigated how these two kinds of PT are reflected in children's implicit and explicit responses.

Social-cognitive skills predicting PT

Gaining the ability to reliably engage in PT likely depends on the other cognitive and social skills that are also developing during early childhood. As children analyze perspectives and act on this analysis, several skills may be relevant for PT: After verbally processing the request, children need to identify the referent according to their partner's perspective, maintain that information in mind, and select the target object. Thus, a second goal of our study was to explore skills that are related to PT. Research on social predictors of PT has found that goal prediction ability (Krogh-Jespersen et al., 2015) and imitation propensity (Herold & Akhtar, 2008) were associated with PT in infants and toddlers. In addition, socio-environmental features, including multilingual exposure (Lieberman et al., 2017), have been related to PT. These social skills and experiences may enhance PT in young children because they promote consideration of one's social partner. In terms of cognitive predictors, vocabulary and working memory (Wardlow, 2013) may enhance reasoning about others' perspectives because children need to process their partner's instructions and maintain another perspective in mind. Importantly, inhibitory control may also affect PT: Because children need to override their own perspectives to instead think about another's view (Baillargeon et al., 2010), inhibition might enhance children's ability to reason about what others know and believe (Carlson & Moses, 2001). In particular, conflict inhibition, or the ability to suppress a prepotent response and deploy a correct response, has been related to young children's awareness of what others can see and may therefore be specifically relevant to PT (Nilsen & Graham, 2009, 2012).

In addition to conflict inhibition, another type of inhibition that may be influential for PT development is inhibition of imitation, which is a skill that bridges the social and cognitive domains. Inhibition of imitation is the ability to withhold automatic tendencies to copy another person's actions. Under certain conditions, such as when perceiving an observed action as intentional (Liepelt, Cramon, & Brass, 2008), adults show the automatic tendency to imitate the action they observe (see Brass, Ruby, & Spengler, 2009 for a review). This tendency is associated with overlapping neural regions that are activated both when performing actions and observing others' actions (e.g., Rizzolatti & Fogassi, 2014) and is evident in mimicry research in children and adults (Chartrand & Bargh, 1999; Van Schaik & Hunnius, 2016). Moreover, patients with lesions in the prefrontal cortex automatically imitated others, lacking the ability to inhibit imitating (Brass, Derrfuss, Matthes-von Cramon, & Von Cramon, 2003).

In certain situations, automatic imitation needs to be inhibited. Across a range of communicative interactions, children may need to suppress spontaneous imitation to maintain the flow of the interaction. For example, turn-taking games involve withholding imitation of a social partner until it is the child's turn to act. Inhibition of imitation has been studied in adults using a finger lifting game. When adults were asked to lift their index

finger but were shown the incongruent action of lifting a middle finger, adults significantly more often performed the incongruent action (middle finger) than the instructed action (index finger; e.g., Brass, Derrfuss, & Von Cramon, 2005), thereby failing to inhibit imitating.

Inhibition of imitation reflects both the cognitive component of inhibition and the social consideration of another's actions (Brass et al., 2009). Inhibiting one's own actions while observing others' actions suppresses the overlap between self and other, increasing the "self-other distinction." The ability to separate oneself from others could benefit PT because PT requires inhibiting one's egocentric perspective to respond appropriately to another's point of view. Indeed, inhibition of imitation has been related to neural areas implicated in PT in adults (tempo-parietal junction; Brass et al., 2005). Further, adults trained to inhibit imitating another's actions performed more accurately in a PT task than those trained to generally inhibit or to imitate others (Santiesteban et al., 2012). It is unknown whether those with better self-other distinction abilities (indexed by inhibition of imitation) have enhanced PT skills, or whether better self-other distinction causes both greater ability to inhibit imitation and engage in PT. Studies have yet to test whether children with enhanced ability to withhold imitation would similarly show better PT skills.

While withholding imitation might benefit PT, it is also possible that a stronger propensity to associate with others would instead bolster PT. Those poorer at withholding automatic imitation tendencies may show increased proneness to consider others and their actions; indeed, toddlers with better imitation abilities were more successful in PT (Herold & Akhtar, 2008). Inhibition of imitation could relate to children's PT in one of two ways: Better ability to withhold imitation could increase the self-other distinction and improve PT (Santiesteban et al., 2012), or poorer ability to withhold imitation could reflect enhanced consideration of others and improve PT. Thus, inhibition of imitation is ripe for study in relation to children's PT as PT necessitates both consideration of one's social partner and the ability to distinguish one's own actions from those of their partner.

Present study

The current study examined how 3-year-olds analyzed a social partner's perspective as they considered and acted in response to their partner's point of view. We chose to study children at an age when we expected they would show variability in their PT performance to evaluate variations in skill across PT responses and situations, and in the hopes of identifying socio-cognitive skills that covary with PT. We designed a task that measured the same PT constructs as in prior work (e.g., Moll & Tomasello, 2006), though the task was not intended to replicate previous studies. Instead, we implemented novel features in the PT task to test children's understanding of what a social partner *can see* and *cannot see*. Further, we examined how PT unfolds by measuring children's initial analysis of their partner's perspective (implicitly via eye gaze) and responses to that analysis (explicit toy choice) within the same task, in contrast to prior research with older children (e.g., Nilsen & Graham, 2009). Importantly, the study and analysis plan were preregistered (<https://aspredicted.org/sz5qx.pdf>). We hypothesized that children would show sensitivity to another's perspective at this age, as indexed by their implicit responses (longer looking to target objects), but might still be challenged to integrate that understanding robustly into their actions, as reflected in their explicit responses (target toy choice). We further hypothesized

that children would perform better when they considered objects their partner could see when compared with objects their partner could not see.

Further, we tested socio-cognitive skills that may support PT. Following Santiesteban et al. (2012), we examined whether inhibition of imitation related to PT, as it could reflect enhanced self-other distinction. We therefore hypothesized that children with greater ability to inhibit imitating would have improved PT abilities. It is also possible that poorer ability to withhold automatic imitation tendencies would relate to PT because it reflects a greater propensity to consider another person. In addition to inhibition of imitation, we tested inhibitory control, which might benefit suppression of children's own perspective or general task performance ability; vocabulary knowledge, which could enhance understanding of the partner's verbal prompt; and working memory, which could allow children to hold the target object in mind as they made a choice. In sum, this study investigated the socio-cognitive factors supporting children's unfolding sensitivity to a social partner's perspective on jointly and uniquely perceived objects in a communicative exchange.

Method

Participants

Thirty-six three-year-olds ($M_{\text{age}} = 37.3$ months; range = 36.0– 38.9 months; 17 female) participated in the study. The sample size was preregistered (<https://aspredicted.org/sz5qx.pdf>). Prior to participation, caregivers provided informed consent and completed a background questionnaire. The majority of parents reported their child was of European descent (25; 3 African American, 1 Asian American, 1 Hispanic or Latin American, 4 two or more races, 2 did not report). Half of the children had caregivers with high levels of education (maternal education: 20 postgraduate, 9 bachelor's, 5 some college, 2 did not report). All children were exposed to English at least 75% of the time at home, were born within three weeks of their due dates, and had no known developmental delays. Thirteen additional children participated in the study but were excluded from analyses due to failing to pass Catch Trials after PT blocks (6; preregistered exclusion), failing to complete socio-cognitive tasks (Inhibition of Imitation or Go-Nogo: 6; preregistered exclusion), or handing both toys on all PT trials (i.e., no usable explicit behavior; 1; see Procedure, Participant Task Inclusion, and Analysis for details on inclusion criteria). The study was conducted in the laboratory of a large research university and took approximately one hour to complete all tasks.

Procedure

Children participated in a PT task and a series of socio-cognitive tasks: Inhibition of Imitation, General Inhibition (Go-Nogo), Vocabulary knowledge (Toolbox Picture Vocabulary Test (TPVT)), and Working Memory (Spin the Pots). All children did the PT task first. The socio-cognitive tasks were presented in one of two fixed orders (18 children per order) to counterbalance the order of the two inhibition tasks (Inhibition of Imitation and Go-Nogo): A) Inhibition of Imitation, TPVT, Go-Nogo, Spin the Pots, or B) Go-Nogo, TPVT, Inhibition of Imitation, and Spin the Pots. The PT task was administered by two experimenters (E1 and E2) and was audio and video recorded simultaneously by four

webcams. Socio-cognitive tasks were administered immediately following the PT task in a separate testing room at a child-sized table by E1, also recorded simultaneously by three webcams.

Perspective-taking task

The PT task was a game in which the child collected toys for animal friends by handing toys to E1 (see [Figure 1](#)) and included PT trials and Control trials. The child could always see both objects, while E1's view of one object was occluded. A verbal prompt indicated which of the two toys the child should hand E1. The prompts either directed the child to consider E1's perspective or use a color control rule to select the target toy. The PT task had four phases: *Familiarization with Experimenter Perspective*, *Introduction*, *Perspective Practice*, and *Test Trials*.

Setup and stimuli

During the PT task, the child sat across the table from E1 ([Figure 1](#)). In front of the child were two colored mats on the child's left and right side (red: child's right, yellow: child's left). During the PT task, objects were placed onto the mats; the child could always see both objects. E1 was hidden from the child's view by an apparatus with two doors. When E1 opened each door (on her left and right), she could see one of two colored mats. In this way, E1's perspective was different from that of the child; with one door open, E1 could see one mat, while the child could always see both. Children were prevented from reaching for the toys on the mats prematurely by a clear plastic barrier that E2 lowered when it was the child's turn to select a toy.

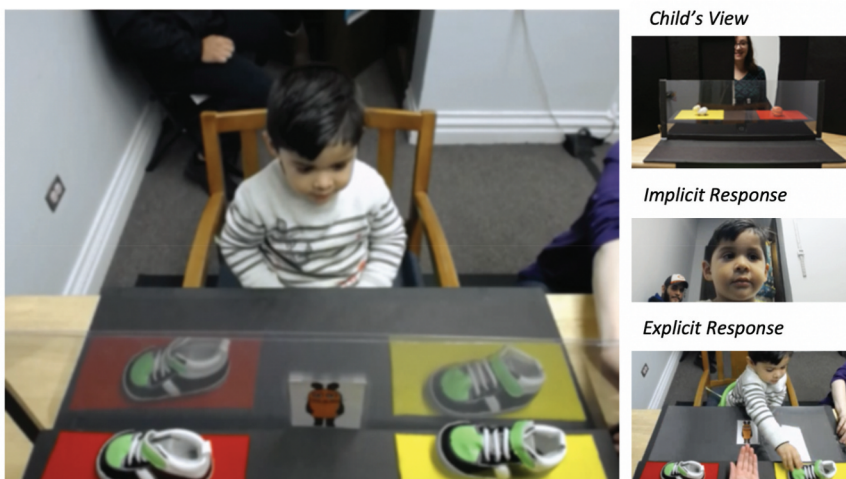


Figure 1. Perspective-taking (PT) task setup.

Note: On each trial, toys were placed on the red and yellow mats. Experimenter 1 (E1) could open one of two doors to view one toy while her view of the other toy was blocked. Children had visual access to both toys. The animal friend in the cover story shared the child's perspective on E1 and the toys. After hearing a prompt about which toy to select, children's eye gaze was coded for two seconds (implicit response). Experimenter 2 (E2) then lowered the barrier, allowing the child to choose a toy (explicit response).

The task had a cover story: Animal friends needed help finding their toys, and it was the child's task to hand the toys to E1. Images of four animal friends (Mouse, Duck, Elephant, and Rabbit) were each attached to the clear barrier, appearing as though the animal shared the child's perspective on the toys and E1. Voiceover of the "animals" indicated which toy the child should choose on each trial. Audio was recorded on Audacity in child-friendly speech and pitch was increased to induce the impression of a cartoon character. E1 controlled the playback of the prompts by unobtrusively pressing a button to advance an EPrime program.

Sixteen identical toy pairs were used for the PT task during *Test Trials* (see Supplementary Materials). Six additional non-identical toy pairs were also used, one pair each for *Familiarization with Experimenter Perspective*, *Perspective Practice*, and each of four Catch Trials. All items were chosen to be known by at least 80% of 30-month-olds (Frank et al., 2016¹).

Procedure

The *Familiarization with Experimenter Perspective* phase introduced the child to E1's perspective on the toys. The child was shown the apparatus from E1's point of view. E1 narrated while the child saw that when each door was opened, one toy at a time could be viewed. For example, E1 opened the door on the right and said, "Now we see a cup! But we don't see anything else." This was repeated for the door on the left. Then the child returned to his or her side of the table for the remainder of the study.

In the *Introduction* phase, E1 explained the task cover story and allowed the child familiarity with the voiceover animal prompts. E1 introduced herself to the first animal friend, invited the child to do the same, and encouraged the child to help the animal find his toys. E1 then affixed the animal card to the barrier. To contrast E1's perspective with that of the child, the *Perspective Practice* phase allowed the child to experience the setup from his or her own perspective with a procedure analogous to *Familiarization with Experimenter Perspective*. Audio prompts narrated the child's and animal's shared perspectives to clarify that the child's view differed from E1's: The child and animal friend can see both toys at all times while E1, with only one door open, can see one toy at a time. For example: With both doors closed, E1 said, "Hmm, I can't see anything over there!" [*E1 opened the door on her right*] "I see a sheep! But I don't see anything else." Mouse voiceover: "We see a sheep and a ball." E1: "Oh really, Mouse? From over here, I just see a sheep. I don't see anything else." This exchange was based on conversations with children during piloting and clarified that the child and animal shared the same perspective, while E1's perspective differed.

Test Trials. Children then participated in as many as four blocks of Test Trials with eight trials per block in the same fixed order across participants. Blocks 1 and 3 were PT blocks, and 2 and 4 were Control trial blocks. PT trials were of two types: Can See and Does Not See. These trials were pseudo-randomized within Block 1 and Block 3, with four trials of each type per block. Control trials were also of two types: Red and Yellow, also counter-balanced and pseudo-randomized with four trials per type. Within each block, the open door side and the correct response side (the child's left or right) were counterbalanced. The same toys used in PT blocks were used in Control blocks.

¹The oldest children in the Wordbank normed sample are 30-month-olds.

On Can See trials, E1 and the child shared visual access to the requested toy; to answer correctly, the child needed to select the toy that could be seen by both people. On Does Not See trials, E1 and the child did not share visual access to the requested toy. To be successful on these trials, children needed to select the toy E1 cannot see, although the child could see both toys. (It should be noted that the child's perspective always differed from that of E1; their respective views on the requested toy differed between Can See and Does Not See trials.) Control trials were implemented because they did not require PT, but children still used a rule to choose one of two toys. The requested toy corresponded to the colored mat on which the toy sat (red or yellow; these are color words children know by this age: Frank et al., 2016). Between each block was a Catch Trial designed to ensure the child was still paying attention (as preregistered).

At the beginning of each Test Trial, E1 had both doors closed. E2 simultaneously placed two identical toys in front of the child, one on each mat. Then, E1 opened one door to give her visual access to one toy. She greeted the child and played the audio prompt in which the animal requested one of the two toys. First, the voiceover specified that a toy was needed: "I need my [toy]." In Can See trials, children were then told, "It's the one [E1 name] can see!" and in Does Not See trials, children heard, "It's the one [E1 name] does not see!" During Control trials, children heard, "It's on the red side!" or "It's on the yellow side!" depending on the trial type.

E1 waited two seconds following the end of the audio prompt, during which children's eye gaze (implicit response) was later coded via a webcam mounted under the platform (see Coding). E1 then extended her hand, saying, "Can I have it?" while E2 lowered the barrier so the child could choose a toy. After the child handed a toy, E1 placed the toy in the box behind her. This toy choice was later coded as the child's explicit response. If the child handed both toys, E1 indicated that one toy should be given. She encouraged reluctant children to hand her a toy and replayed the audio prompt if necessary until the child handed a toy.

For Catch Trials between blocks, E2 placed two different toys on the mats. E1 opened both doors and requested one of two non-identical toys by name. These trials ensured no PT or color knowledge was required for a correct response. If a child did not correctly respond to the Catch Trial, data from the prior block were excluded (as preregistered). To be included in analyses, children needed to complete at least one PT block followed by a correct Catch Trial.

Socio-cognitive tasks

Inhibition of Imitation

This task was designed to measure children's ability to inhibit the tendency to imitate another person's actions. This Stroop-like task was adapted for use with children (based on Brass et al., 2009). Children were instructed to press one of two round, colored buttons in response to a colored image presented on a computer screen. Images showed a person from torso to neck, with one hand pressing the yellow (left), green (right), or neither of two large buttons. Images were completely shaded either yellow or green (see Figure 2). Children were given a corresponding yellow and green button and were instructed to press the green button when viewing a green image and the yellow button when viewing

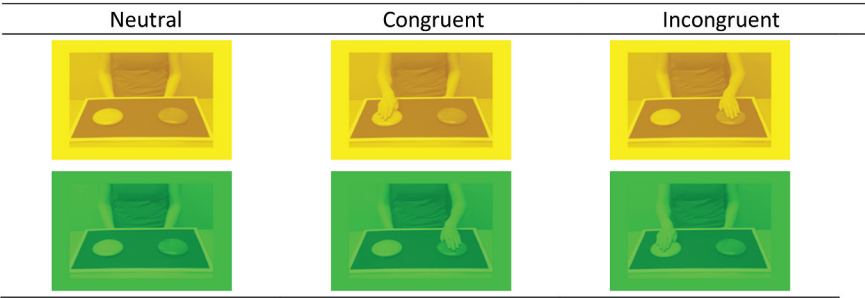


Figure 2. Inhibition of imitation stimuli.
Note: Children were instructed to press a corresponding yellow button when viewing yellow pictures, and a green button when viewing green pictures. Neutral trials do not contain any action information. Congruent trials depict an action consistent with the action the child should perform to be successful. Incongruent trials depict an action opposite to the action the child should perform to be successful. Accuracy on each trial type was measured.

a yellow image. Notably, the color of the image and the button pressed in the image were crossed to generate three trial types: Congruent, in which the color of the image matched the button the person in the image was pressing; Incongruent, in which the color of the image was opposed to the button being pressed in the image; and Neutral, in which there was no competing action information. Children’s buttons mirrored the buttons in the images, thereby increasing children’s tendency to copy the actions they saw. However, to be successful, children needed to ignore the actions (i.e., withhold imitation) and follow the color rule. If children instead followed the actions of the person, they would respond incorrectly on the Incongruent trials. Therefore, worse performance on Incongruent compared to Congruent trials reflects automatic imitation tendencies that were not suppressed. While we expected that as a group children would demonstrate a “Congruency Effect” where performance would be greater on Congruent than Incongruent trials, we also expected individual differences in how well children could inhibit their automatic imitation tendencies.

E1 first trained children to name the color of the pictures and press the button corresponding with the colored image using Neutral trials. Children received corrective feedback on one Congruent and one Incongruent trial of each color. Children then participated in two blocks of eighteen test trials each, with six trials of each type (Congruent, Incongruent, and Neutral) per block, half yellow and half green, in a fixed pseudo-randomized order. Each image appeared on the screen with a simultaneous sound that was identical for all trials. After training and between trial blocks, a happy face appeared accompanied by a sound and E1 reminded the child of the game rules. To be included in analyses, children completed at least one of the two test blocks.

Go-Nogo

A classic inhibitory control measure was adapted for use with children. This task assessed response inhibition: children responded to one stimulus and withheld a response to another. Children were presented with a red button centered in front of a computer screen.

Images presented on the screen were blue shapes (heart and star) on white backgrounds. Each image was displayed for four seconds or until the child pressed the button. E1 trained children with four images: Children identified the shapes and were shown how to press the button when viewing the star (“Go” trials) and “just wait” when viewing the heart (“Nogo” trials). Children then received corrective feedback on four additional trials, two of each shape. Finally, children completed test trials in two blocks of twelve trials each, with six trials of each type in the same fixed pseudo-randomized order for all children. Analogous to the Inhibition of Imitation task, the same sound played when each image appeared on the screen. Between blocks, a smiley face appeared on the screen and E1 reminded children of the game rules. Children included in analyses completed at least one block of test trials.

TPVT

Children’s vocabulary knowledge was assessed with the Toolbox Picture Vocabulary Test (TPVT) on the NIH Toolbox application on an iPad (Gershon et al., 2010). This normed, adaptive test prompted children to touch one of four photographs that most closely matched the word they heard.

Spin the Pots

Children’s working memory was assessed with the Spin the Pots task (Hughes & Ensor, 2005; adapted for three-year-olds in Beck et al., 2011). Children were instructed to find eight stickers, each hidden in one of ten boxes (with two boxes empty). On each trial, the boxes were covered with a cloth and spun 360-degrees. E1 removed the cloth and asked, “Which box should we open?” E1 opened the first box the child pointed to or touched. Children were given a maximum of 16 trials to find all stickers.

Coding

PT task coding

All coding was done off-line by two coders in Interact (Mangold, 2017) blind to study hypotheses. Children’s *explicit response*, i.e., the toy handed to E1, was coded as a correct (target) or incorrect (distractor) toy choice relative to the trial type. Invalid responses consisted of giving neither or both toys, or by moving the toys around so it was no longer clear from which mat each toy originated. Additionally, children’s *implicit response* was coded by measuring eye gaze toward the target and distractor toy via a webcam mounted below the platform. Eye gaze was coded toward the target and distractor toy from the frame after the end of the prompt (e.g., “It’s the one [E1 name] can see”) until the frame before the barrier was lowered (approximately two seconds), thereby separating the implicit response window from explicit responses in time. Gaze was coded as visual fixations of at least two frames at a rate of 30 frames per second.

Socio-cognitive task coding

Children's button presses on the Inhibition of Imitation and Go-Nogo task were recorded by an EPrime program. Analyses used measures of accuracy for Inhibition of Imitation and Go-Nogo because reaction time was extremely variable (see Supplemental Materials). Inhibition of Imitation accuracy was calculated by summing the number of correct button presses on Congruent, Incongruent, and Neutral trials following the color rule. Go-Nogo accuracy was calculated by summing the number of Go trials on which the child pressed the button and the number of Nogo trials on which the child did not press the button. In addition, a research assistant coded children's visual attention during both tasks to exclude trials on which children were not looking at the screen prior to making or withholding a response.

The TPVT was scored automatically and age-normed scores were used for analyses. Spin the Pots was coded following Hughes and Ensor (2005): Scores were the number of errors children made on the task subtracted from the total number of possible spins (16). Errors were spins on which children selected boxes that did not contain a sticker either because that box was empty from the start or because the child had already opened that box and removed a sticker.

Analysis

Reliability

Eight of the 36 children (22.2%) were double-coded for reliability on all measures by an independent coder blind to hypotheses. Reliability on explicit and implicit PT responses was high (explicit: proportion target toy choice: Cronbach's $\alpha = 0.999$; ICC = 0.999, $F(7) = 836.5$, $p < .001$; implicit: proportion of time looking to the target object out of total time looking to target and distractor object: Cronbach's $\alpha = 0.978$; ICC = 0.969, $F(7) = 46.4$, $p < .001$). Reliability on trials excluded from Inhibition of Imitation and Go-Nogo was also high (measured with Cohen's kappa (κ): Inhibition of Imitation: κ could not be calculated (no trials were excluded for one coder): 98.4% agreement on trial exclusion; Go-Nogo: $\kappa = 0.664$, $p < .001$). Reliability on Spin the Pots score was also high: $\kappa = 1.000$, $p < .001$.

Participant task inclusion

All children included in analyses completed at least one PT block and 34 of 36 children also completed one or more Control blocks (see Perspective-Taking Task). For two children, Inhibition of Imitation and Go-Nogo were not video recorded due to equipment failure and could therefore not be coded for visual attention (see Coding). Despite potentially contributing noise in the data, trials from these two children were included in the analyses. One child did not have a TPVT score due to technical failure. Three children lacked scores for the Spin the Pots task (video file was corrupted (1), child refused to complete the task (1), experimenter error (1)).

Perspective-taking task

PT and Control explicit responses were converted into proportions by dividing the number of correct trials by the total number of trials the child completed, which yielded a proportion out of 1. For implicit responses, the amount of time children spent looking at the target object was divided by the total time spent looking at the target and distractor object ($\text{target}/(\text{target} + \text{distractor})$) on each trial to yield a proportion out of 1 for each trial. Proportions were then averaged across trials. Longer looking toward the target toy (compared to the distractor toy) was conceptualized to index children's ability to differentiate between the requested and non-requested toy based on the PT or Control prompt. This is analogous to preferential looking time studies in infancy, which reflect infants' ability to discriminate between two stimuli (e.g., Aslin, 2007).

Socio-cognitive tasks

For Inhibition of Imitation and Go-Nogo, accuracy per trial type was used to calculate summary composite scores (Congruency Score and Inhibition Score). The Inhibition of Imitation Congruency Score reflected both better performance on Congruent compared to Incongruent trials and children's overall performance across both trial types: $(\text{Congruent} - \text{Incongruent}) * ((\text{Congruent} + \text{Incongruent})/2)$. Positive scores reflect better accuracy on Congruent than Incongruent trials accounting for overall accuracy. We expected that children would show a "Congruency Effect," performing more accurately on Congruent than Incongruent trials overall. We then planned to analyze whether the extent to which children were affected by another's actions was related to PT using the Congruency Score.

However, children as a group did not demonstrate a Congruency Effect: only 13 children showed the overall effect. Thus, for exploratory purposes, children were split by whether or not they showed the Congruency Effect. This split was performed based on raw accuracy data from Congruent and Incongruent trials: Those who were more accurate on Congruent than Incongruent trials were categorized as demonstrating a Congruency Effect (13 children), in contrast to those who did not show this effect (23 children). Those in the Congruency Effect group showed automatic imitation tendencies and were affected by the actions they viewed. The group that did not demonstrate the Congruency Effect was more heterogeneous in performance and did not demonstrate the expected behavioral pattern of poorer performance when action information competed with the color task.

An Inhibition Score was calculated for Go-Nogo similarly to the Congruency Score: $\text{Nogo} * ((\text{Go} + \text{Nogo})/2)$. Higher scores indicate more inhibition while accounting for performance on Go trials. An exploratory data split was also performed, creating groups of children who did not inhibit versus those who did inhibit. Children who scored less than 50% on Nogo trials were categorized as Non-Inhibition (17 children) and those who scored greater than 50% were categorized as Successful Inhibition (19 children). All data were examined for outliers (more than 3 standard deviations from the mean; no outliers were found). Proportion data were arcsine-square-root transformed for analyses.

Results

We first report results of children's performance on the PT task, in the different trial types and both explicitly and implicitly. We then tested whether PT performance was related to children's socio-cognitive skills. Deviations from preregistered analyses are mentioned where applicable (see Supplementary Material for all planned analyses).

PT and Control performance

First, we examined explicit and implicit performance on PT and Control trials (Figure 3). On Control trials, children reliably handed and looked longer to the requested toy compared to chance performance (explicit: $t(33) = 5.05$, $p < .001$, $d = 0.866$; implicit: $t(33) = 7.66$, $p < .001$, $d = 1.313$). For PT trials, children did not consistently choose the target toy; performance was not significantly above chance for explicit PT performance ($t(35) = 1.64$, $p = .111$). However, children did look significantly longer to the target than distractor toy on PT trials (compared to chance: $t(34) = 5.94$, $p < .001$, $d = 0.702$). Therefore, children's implicit responses indicated sensitivity to another's perspective, but their explicit performance did not reflect accounting for another's perspective when acting.

PT trial types (explicit and implicit)

Next, we tested whether explicit and implicit PT performance differed when the requested object was jointly or individually perceived (Can See and Does Not See trials; Figure 4). Children's explicit behavior indicated above-chance performance on Can See trials ($t(35) = 2.70$, $p = .011$, $d = 0.449$), while children did not reliably choose the target toy on Does Not See trials (performance did not differ from chance: $t(35) = -0.29$, $p = .772$). Children more often selected the target object on Can See than Does Not See trials ($t(35) = 2.22$, $p = .033$, $d = 0.519$). Implicitly, children looked longer to the target object compared to chance on both Can See ($t(35) = 6.10$, $p < .001$, $d = 1.106$) and Does Not See ($t(34) = 2.55$, $p = .015$, $d =$

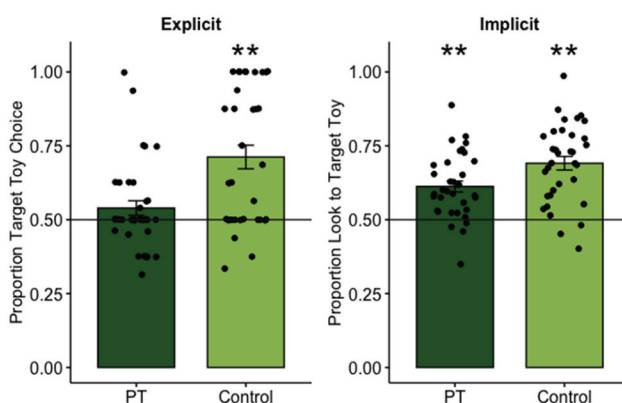


Figure 3. Perspective-taking (PT) and control trial performance.

Note: Average explicit response (proportion target toy choice) and average implicit response (proportion time looking to target toy) on perspective taking (PT) and Control trials. Chance is .50. Error bars are ± 1 Standard Error. ** $p < .001$.

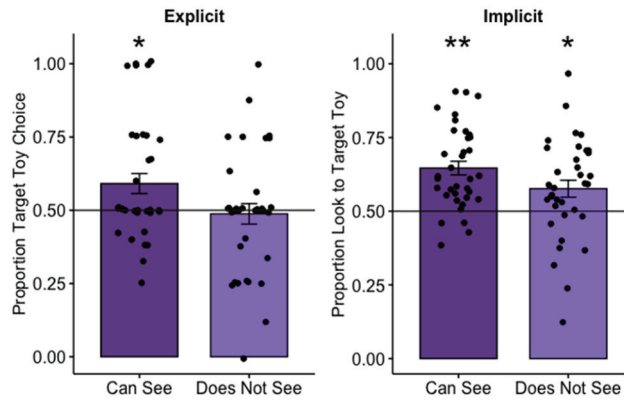


Figure 4. Perspective-taking (PT) performance by trial type.

Note: Performance on Can See and Does Not See trials, measured explicitly and implicitly. Chance is .50. Error bars are ± 1 Standard Error. ** $p < .001$, * $p < .05$.

0.431) trials (implicit responses did not differ significantly on Can See and Does Not See trials: $t(34) = 1.86$, $p = .072$). These analyses suggest that in both object perception situations, children's implicit responses indicated sensitivity to another's perspective, but children were not consistently able to implement PT sensitivity into their explicit responses on the Does Not See trials.

Exploratorily, we examined whether children's initial eye gaze toward the target toy and subsequent toy choice were related to one another on PT trials. Explicit and implicit responses were significantly correlated ($r = 0.342$, $p = .044$; see Figure 5). However, explicit

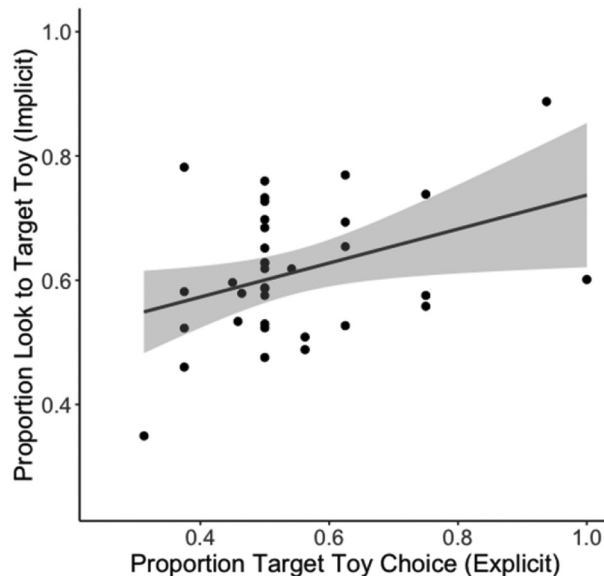


Figure 5. Perspective-taking (PT) task explicit and implicit performance.

Note: Scatterplot showing the relation between explicit and implicit measures of PT.

Table 1. Perspective-taking (PT) task McNemar test: explicit and implicit performance.

	Implicit Response (< .50 looking to target toy)	Implicit Response (> .50 looking to target toy)
Explicit Response (< .50 target toy choice)	3	20
Explicit Response (> .50 target toy choice)	1	11

Table of McNemar test results showing the number of children who demonstrated selection of the target toy explicitly (> .50 selection of target toy) and implicit gaze toward the target toy rather than the distractor (> .50 looking to target toy).

and implicit PT responses were also distinct; as shown previously, performance was above chance for explicit Can See trials and not different from chance for Does Not See trials, while performance was above chance for both trial types when measured implicitly. We ran an exploratory McNemar Test (Table 1) examining whether children who were more likely to respond accurately in their behavior (greater than .50 target toy choices) would also be more likely to look at the requested object (greater than .50 looking to the target toy). This test was significant ($p < .001$): While 11 children demonstrated systematic explicit responses and longer looking to the target object, 20 children looked longer to the target toys but did not choose the correct objects. Moreover, only one child who failed to look longer to the target object performed correctly when making a choice. Thus, while explicit and implicit performance were related, children demonstrated sensitivity to PT as measured by gaze, but did not consistently integrate this analysis in their behavior. This suggests that children can successfully analyze another's perspective, though explicit performance may be reduced when children execute a choice.

PT and socio-cognitive skills

An overview of the descriptive statistics for each of the socio-cognitive tasks is presented in Table 2. Correlations among these socio-cognitive skills were only evident between general inhibition and working memory (negatively; $p = .012$); the other measures were not correlated (all p 's > .090; see Table 3).

Prior to analyzing socio-cognitive skills in relation to PT, we examined performance on the Inhibition of Imitation task. Since to our knowledge, inhibition of imitation has not been assessed in young children, we tested whether children as a group showed differential

Table 2. Descriptive statistics of socio-cognitive measures.

		Mean	SD	Range
Inhibition of Imitation	Congruent	75.00%	23.15%	20.00–100%
	Incongruent	72.82%	24.52%	20.00–100%
	Neutral	76.16%	21.74%	33.33–100%
General Inhibition	Go	70.09%	35.59%	0–100%
	Nogo	54.12%	42.16%	0–100%
Vocabulary		94.29	11.58	68–112
Working Memory		11.79	3.39	4–16

Table of descriptive statistics for performance on socio-cognitive tasks of Note: Scatterplot showing the relation between explicit and implicit measures of PT., Inhibition of Imitation, vocabulary, and working memory. Values indicate average accuracy: Percentage correct on each Inhibition of Imitation task trial type, percentage correct for each General Inhibition task trial type, Vocabulary score on the Toolbox Picture Vocabulary Test (TPVT), and Working Memory (Spin the Pots) score. Standard deviations and ranges are listed.

Table 3. Socio-cognitive measures correlations.

		Inhibition of Imitation	General Inhibition	Vocabulary	Working Memory
Inhibition of Imitation					
	General Inhibition				
	Pearson's <i>r</i>	−0.052			
	Significance	.761			
Vocabulary					
	Pearson's <i>r</i>	0.007	0.291		
	Significance	.970	.090		
Working Memory					
	Pearson's <i>r</i>	0.187	−0.425*	0.011	
	Significance	.291	.012	.953	

Table of bivariate correlations between socio-cognitive tasks of Inhibition of Imitation, Inhibition of Imitation, vocabulary, and working memory. Values indicate Person's *r* and two-sided *p*-values for significance. **p* < .05.

accuracy on Congruent, Incongruent, and Neutral trials of Inhibition of Imitation. A preregistered repeated-measures ANOVA found no significant difference between children's accuracy on any trial type ($F(2,34) = 0.118$, $p = .889$, $\eta_p^2 = 0.007$). This contrasts with adult reaction time findings (faster responses for Congruent than Incongruent trials), which suggested that children's group-level performance would be more accurate on Congruent than Incongruent trials. Since the Inhibition Score relied on the assumption that children as a group would show a Congruency Effect and that variations in the strength of this effect might relate to PT, this raised concerns. Therefore, rather than relying on variations in Inhibition Score, we analyzed group differences by exploratorily splitting children by those who were affected by another person's actions (i.e., showing the Congruency Effect: 13 children) and those who were not affected (23 children). Analogously, we split children by their general inhibition (Go-Nogo) performance: The Successful Inhibition group (19 children) demonstrated more than 50% inhibition on Nogo trials, while the Non-Inhibition group (17 children) showed less than 50% inhibition on Nogo trials. These groups were used in the next exploratory analyses (see Supplemental Materials for pre-registered analyses).

To test whether children's socio-cognitive skills were related to PT, two exploratory repeated-measures ANOVAs were run on Can See and Does Not See performance, one ANOVA for explicit performance and one for implicit performance. In both models, all socio-cognitive measures were included: Inhibition of Imitation Congruency Groups and Go-Nogo Successful Inhibition and Non-Inhibition groups were between-subjects factors, and TPVT and Spin the Pots were covariates. In contrast to initial preregistered analyses, this parsimonious strategy allowed us to test the effect of each socio-cognitive skill on implicit and explicit responses, for Can See and Does Not See trials, in two models. Only Inhibition of Imitation (measured by Congruency Group) showed a significant interaction with explicit performance on Can See and Does Not See trials: $F(1,27) = 4.371$, $p = .046$, $\eta_p^2 = .139$; all other main effects and interactions were not significant (p 's > .514). While inhibition of imitation did not significantly interact with implicit performance ($F(1,26) = 0.286$, $p = .609$, $\eta_p^2 = .010$), results trended in the same direction as in the analysis of explicit responses (see Figure 6). The analysis of explicit responses suggested that children demonstrating a Congruency Effect did not differ in performance on Can See and Does Not See trials ($t(12) = -0.719$, $p = .486$) while those not demonstrating this effect performed more accurately on Can See than Does Not See trials ($t(22) = 2.811$, $p = .010$, $d = 0.834$; see Figure 6). Further, children demonstrating a Congruency Effect showed increased performance on Does Not See trials compared to those who did not demonstrate the effect ($t(34) = 2.133$, $p =$

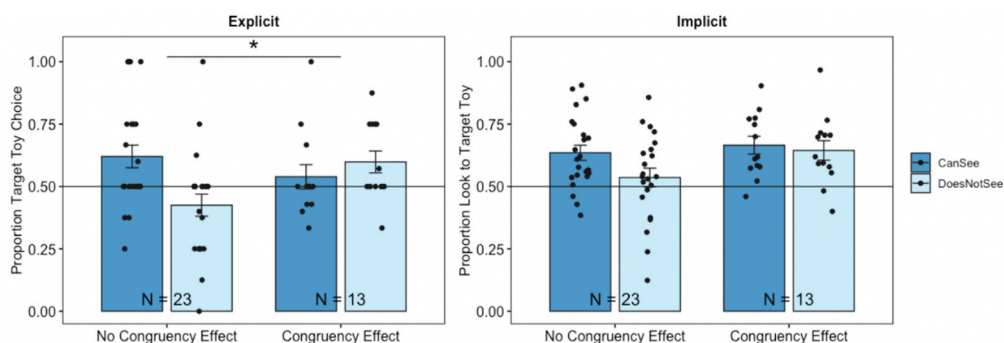


Figure 6. Perspective-taking (PT) performance by inhibition of imitation groups and trial type. Note: Explicit and implicit performance on Can See and Does Not See trials, split by whether children showed a Congruency Effect in Inhibition of Imitation performance. 13 children showed the Congruency Effect and 23 children did not. Chance is .50. Error bars are ± 1 Standard Error. * $p < .05$.

.040, $d = 0.790$). Therefore, children who showed an automatic tendency to imitate another person without being able to inhibit this tendency (i.e., those more affected by another's actions) demonstrated better explicit PT performance, particularly on the Does Not See trials. In sum, inhibition of imitation performance related to differential accuracy on Can See and Does Not See trials, while the socio-cognitive predictors of inhibition, vocabulary, and working memory did not evidence relations with PT.

Of note, there were no indications that children who demonstrated versus did not demonstrate a Congruency Effect differed in other aspects of performance. We did not see evidence that children in the two groups significantly differed in their Neutral trial accuracy ($t(12) = 0.400$, $p = .696$; see Supplementary Materials), vocabularies (TPVT: $t(11) = 0.164$, $p = .873$), ages ($t(12) = 1.601$, $p = .135$), or general inhibition skills (Nogo accuracy: $t(12) = 0.309$, $p = .763$). Thus, it appears that children who showed versus did not show a Congruency Effect differed only in their Incongruent versus Congruent trial performance, suggesting that the PT performance difference seen between groups is likely due to inhibiting imitation, rather than other child-level characteristics.

Results summary

When first analyzing another person's perspective, children performed above-chance on both PT trial types as measured implicitly by their gaze. When children subsequently acted on their knowledge, they demonstrated greater explicit PT performance on Can See than Does Not See trials; performance was above chance on Can See trials and at chance on Does Not See trials. Additionally, children's PT skills differed depending on their ability to inhibit imitating. Contrary to adult findings, children who showed a Congruency Effect, i.e., those who struggled to inhibit imitating, performed better on the PT task, specifically on the Does Not See trials. Children more affected by another person's actions were better able to consider another's perspective. Importantly, this relation appeared specific to inhibiting imitating; we did not see evidence that children's ability to generally inhibit, vocabulary knowledge, or working memory capacity was related to differential PT performance.

Discussion

In this study, we tested socio-cognitive skills supporting 3-year-olds' sensitivity to and subsequent ability to act on another's perspective in a social interaction. As children analyzed their partner's perspective, their gaze was measured implicitly as an early indicator of sensitivity to another's perspective. This analysis process culminated in children's toy choice, indexing a later phase of PT-based behavior that tested whether children could act according to their PT analysis. PT was assessed when children considered whether a social partner *could see* and *could not see* a requested toy that was always visible to the child. Further, socio-cognitive skills of inhibition of imitation, general inhibition, vocabulary, and working memory were measured and tested in relation to children's PT responses. Based on findings in adults, we hypothesized inhibition of imitation skills would predict children's PT skills beyond general inhibition.

PT in social interactions

To successfully contribute to social interactions, children need to use their knowledge about another's perspective and integrate this analysis in their behavior. Here, children's gaze behavior suggested that they accounted for their partner's perspective when a requested object was and was not jointly perceived, yet children did not consistently implement PT in their choice behavior. Children's explicit responses indicated that they more accurately accounted for their partner's perspective in situations where both the child and their partner had visual access to a requested object, while performance was inconsistent when the requested object was outside of their partner's view. Implicit and explicit responses were correlated in both object perception situations, yet distinct when children considered objects outside of their partner's view. This suggests that, in line with prior findings (e.g., Surtees & Apperly, 2012), PT may be more difficult to implement when considering objects that are not jointly perceived.

Why did children demonstrate sensitivity to what another could not see when measured implicitly, but fail to consistently translate that analysis into behavior? Perhaps when in doubt, children default to acting on jointly perceived objects in social interactions. The literature on joint attention's central role in social interactions (e.g., Tomasello, 1995) suggests that children attend to objects they and their partner view. The salience of jointly perceived objects may cause children to select this object even when it is not appropriate. However, this object bias was specific to trials in which another's perspective was relevant and was not present in Control trials. This may be because Control trials did not necessitate social engagement; children could respond by looking at the mats under the toys without considering their partner. Thus, though children successfully analyzed their partner's perspective, social interactions may cause children to occasionally preferentially act on jointly perceivable objects. With development, this tendency may be reduced, allowing children to act in accordance with their PT knowledge and suppress the potential salience of jointly viewed objects.

Another explanation for the difference in performance across object perception situations is that when considering what another person *cannot see*, children may need to override an initial assessment of what that person *can see*, potentially decreasing performance when perspectives differ. It is also possible that the linguistic complexity of negation

specific to Does Not See prompts (“It’s the one E1 *does not see*”) may make these trials more difficult. If this were the case, we would then expect children’s general inhibition or vocabulary skills to enhance PT; however, we saw no relations between either skill and PT performance. Furthermore, these factors would have affected children’s implicit performance as well. Since children responded at above-chance levels implicitly, these alternative explanations are unfounded. In sum, similar to findings in older children (e.g., Nilsen et al., 2008), PT *understanding* appears to be online and robust by age three, while *acting* on PT knowledge may still be developing.

PT and inhibition of imitation

In addition to providing evidence for how children consider others’ perspectives early in life, results indicated that specific socio-cognitive skills may be relevant for PT. In adults, training to inhibit spontaneous imitation was related to PT (Santesteban et al., 2012). Successfully inhibiting imitating indicates greater distinction between one’s own actions and those of another, while failing to inhibit imitating (i.e., spontaneously imitating) may reflect social engagement. While adults trained to inhibit imitating had better PT skill, we found evidence for the opposite pattern in children: Those who performed poorly when action information competed with the child’s own action demonstrated better PT skill, particularly when perspectives differed. This was found by exploratorily grouping children by inhibition of imitation performance: Children who failed to withhold spontaneous imitation demonstrated improved explicit PT performance. In other words, those more affected by another’s actions better considered another’s point of view.

Specifically, children who showed the Congruency Effect imitated actions that competed with their instructed actions (performance on Incongruent trials was significantly poorer than Neutral trials; see Supplemental Materials). Children who did not demonstrate this effect did not show differential performance by trial type. Thus, children showing the Congruency Effect were affected by another’s actions; they imitated actions specifically when not directed to do so, while those who did not show the effect less often did so. Thus, children who followed another’s actions and were therefore more sensitive to another’s behavior had improved PT performance. These findings align with research in toddlers showing a relation between imitation fidelity and PT (Herold & Akhtar, 2008). We have extended this finding by incorporating the cognitive component of inhibition, similar to research in older children showing a link between inhibition and PT (Nilsen & Graham, 2009).

Still, this pattern of results was found by isolating a group of children who demonstrated the expected Congruency Effect; it is unknown whether this relation would hold when considering levels of automatic imitation *within* the group of children who spontaneously copied others’ behavior. Since we lacked a large enough group of children who were affected by another’s actions, we could not test to what extent children’s inhibition of imitation related to PT. Rather, the present findings may indicate a threshold children need to reach to consider their social partner *at all* before they can consider their perspective. Perhaps during situations requiring PT, after engaging with the social partner, children (like adults) use inhibitory skills to enhance a self-other distinction to effectively engage in PT. Thus, results presented here could reflect the same mechanisms supporting PT in children and adults, though perhaps tapping the initial social engagement children need to engage in PT.

Studies could follow up on these results to assess the different roles that the propensity to consider others' actions and social distancing play in considering another's perspective.

Interestingly, we did not see relations between general inhibition, vocabulary knowledge, or working memory and PT performance. As the PT decision process is complex and likely involves other socio-cognitive skills beyond social engagement, we expected to see a relation between such skills and PT. Indeed, prior research suggests that working memory and general inhibition, for example, may be beneficial in maintaining information in mind and selecting requested objects over non-requested objects when considering another's point of view (e.g., Wardlow, 2013). Though such relations were not found here, this lack of evidence should be interpreted with caution. It is possible that these skills are at work in promoting PT, but were not fully captured in this study.

PT task contributions

Beyond the results found here, the PT task developed for use in this study is itself beneficial to the field. This task incorporated new elements not used in prior studies, which allowed for a more complete and nuanced measurement of children's PT skills. Here, multiple trials allowed for robust measurements of PT sensitivity, incorporating analysis of jointly versus individually perceived referents, which had not yet been tested in young children (e.g., Krogh-Jespersen et al., 2015). By testing both implicit and explicit PT performance, this task examined children's initial PT analysis and subsequent action based on that analysis within the same task, in contrast to tasks used with older children that tested responses separately (e.g., Nilsen et al., 2008). Further, we interleaved perspective and control trials within a social, communicative task. Future studies could utilize this task to measure PT skills at different ages in development.

Similarly, we adapted an inhibition of imitation measure (Brass et al., 2009) for use with children. While reaction time was used in adults, accuracy may be a better measure of children's skill in this task. It should also be noted that on group level, the Congruency Effect was not replicated in children – the task and the scoring system could therefore be modified for use in subsequent studies to fully capture children's ability to inhibit imitating. The inability to suppress imitating another's actions is an interesting measure as well; it could reflect spontaneous processing of another's actions, mirroring, or mimicry. Developmental work on inhibition of imitation would benefit our understanding of the constructs this task measures and how inhibition of imitation develops through life. Indeed, inhibiting imitation of others' actions may become increasingly difficult across the lifespan in older adults (Wermelinger, Gampe, Behr, & Daum, 2018).

Conclusion

In sum, this study represented a novel approach to elucidate children's unfolding PT analysis and its integration in behavior early in development within the context of communicative, social interactions. As children considered their social partner's perspective, early stages of perspective analysis were measured implicitly via gaze, and later stages were measured via toy choice. This unfolding process reflected both children's sensitivity to another's perspective and their ability to act on that knowledge. By age three, children's gaze indicated that they were sensitive to another's perspective on referents that were jointly perceived and those that were only visible

to the child; still, children struggled to act when referents were not jointly viewed. More work is necessary to determine when in life children can consistently integrate PT sensitivity into action. In addition, children who spontaneously imitated another's actions when they were tasked to withhold imitation showed better understanding of what another person could not see. The propensity to consider others' actions even when not instructed to do so may be a crucial first step in considering another's point of view. Overall, this study contributes to our understanding of children's ability to consider others' perspectives and integrate this understanding in social interactions, and sheds light on the socio-cognitive processes supporting PT early in life.

Acknowledgments

We would like to thank River MacLeod and Kayleigh Watson for their assistance with coding. We would also like to thank the children and families who participated.

Disclosure statement

No potential competing interest was reported by the authors.

Funding

This work was supported in part by the NIHCD grant under [Grant P01 HD064653] and the Institute of Education Sciences, U.S. Department of Education, under [Grant R305B140048] at the University of Chicago. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

ORCID

Natalie Brezack  <http://orcid.org/0000-0001-5236-9721>

Marlene Meyer  <http://orcid.org/0000-0003-2229-6933>

Amanda L. Woodward  <http://orcid.org/0000-0002-9383-9969>

Data availability statement

The data that support the findings of this study are available from the corresponding author, NB, upon reasonable request. Email: nbrezack@uchicago.edu for access to data.

References

- Aslin, R. N. (2007). What's in a look? *Developmental Science*, 10(1), 48–53. doi:10.1111/j.1467-7687.2007.00563.x
- Bahtiyar, S., & Küntay, A. C. (2009). Integration of communicative partner's visual perspective in patterns of referential requests. *Journal of Child Language*, 36(3), 529–555. doi:10.1017/S0305000908009094
- Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, 14(3), 110–118. doi:10.1016/j.tics.2009.12.006
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, 21(1), 37–46. doi:10.1016/0010-0277(85)90022-8

- Beck, D. M., Schaefer, C., Pang, K., & Carlson, S. M. (2011). Executive function in preschool children: Test-retest reliability. *Journal of Cognition and Development*, 12(2), 169–193. doi:10.1080/15248372.2011.563485
- Brass, M., Derrfuss, J., Matthes-von Cramon, G., & Von Cramon, D. Y. (2003). Imitative response tendencies in patients with frontal brain lesions. *Neuropsychology*, 17(2), 265–271. doi:10.1037/0894-4105.17.2.265
- Brass, M., Derrfuss, J., & Von Cramon, D. Y. (2005). The inhibition of imitative and overlearned responses: A functional double dissociation. *Neuropsychologia*, 43(1), 89–98. doi:10.1016/j.neuropsychologia.2004.06.018
- Brass, M., Ruby, P., & Spengler, S. (2009). Inhibition of imitative behaviour and social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1528), 2359–2367. doi:10.1098/rstb.2009.0066
- Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, 112(2), 337–342. doi:10.1016/j.cognition.2009.05.006
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, 72(4), 1032–1053. doi:10.1111/1467-8624.00333
- Caron, A., Kiel, E., Dayton, M., & Butler, S. (2002). Comprehension of the referential intent of looking and pointing between 12 and 15 months. *Journal of Cognition and Development*, 3(4), 445–464. doi:10.1207/S15327647JCD3.4.04
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 76(6), 893. doi:10.1037/0022-3514.76.6.893
- Dunphy-Lelii, S., & Wellman, H. M. (2004). Infants' understanding of occlusion of others' line-of-sight: Implications for an emerging theory of mind. *European Journal of Developmental Psychology*, 1(1), 49–66. doi:10.1080/17405620444000049
- Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, 40, 760–768. doi:10.1016/j.jesp.2004.02.002
- Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2016). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, 44(3), 677. doi:10.1017/S0305000916000209
- Gershon, R. C., Cella, D., Fox, N. A., Havlik, R. J., Hendrie, H. C., & Wagster, M. V. (2010). Assessment of neurological and behavioural function: The NIH toolbox. *The Lancet Neurology*, 9(2), 138–139. doi:10.1016/S1474-4422(09)70335-7
- Herold, K. H., & Akhtar, N. (2008). Imitative learning from a third-party interaction: Relations with self-recognition and perspective taking. *Journal of Experimental Child Psychology*, 101(2), 114–123. doi:10.1016/j.jecp.2008.05.004
- Hughes, C., & Ensor, R. (2005). Executive function and theory of mind in 2 year olds: A family affair? *Developmental Neuropsychology*, 28(2), 645–668. doi:10.1207/s15326942dn2802_5
- Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, 11(1), 32–38. doi:10.1111/1467-9280.00211
- Khu, M., Chambers, C. G., & Graham, S. A. (2020). Preschoolers flexibly shift between speakers' perspectives during real-time language comprehension. *Child Development*, 91(3), e619–e634. doi:10.1111/cdev.13270
- Krogh-Jespersen, S., Liberman, Z., & Woodward, A. L. (2015). Think fast! The relationship between goal prediction speed and social competence in infants. *Developmental Science*, 18(5), 815–823. doi:10.1111/desc.12249
- Liberman, Z., Woodward, A. L., Keysar, B., & Kinzler, K. D. (2017). Exposure to multiple languages enhances communication skills in infancy. *Developmental Science*, 20(1), 1–11. doi:10.1111/desc.12420
- Liepelt, R., Cramon, D., & Brass, M. (2008). What is matched in direct matching? Intention attribution modulates motor priming. *Journal of Experimental Psychology. Human Perception and Performance*, 34(3), 578. doi:10.1037/0096-1523.34.3.578

- Luo, Y., & Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition* 105(3), 489–512. [10.1016/j.cognition.2006.10.007](https://doi.org/10.1016/j.cognition.2006.10.007)
- Mangold (2017): INTERACT User Guide. Mangold International GmbH (Ed.) www.mangold-international.com
- Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E., & Flavell, J. H. (1974). The early development of inferences about the visual percepts of others. *Child Development*, 45(2), 357–366. doi:10.2307/1127956
- McCleery, J. P., Surtees, A. D. R., Graham, K. A., Richards, J. E., & Apperly, I. A. (2011). The neural and cognitive time course of theory of mind. *Journal of Neuroscience*, 31(36), 12849–12854. doi:10.1523/JNEUROSCI.1392-11.2011
- Moll, H., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology*, 24(3), 603–613. doi:10.1348/026151005X55370
- Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, 13(4), 329–336. doi:10.1111/j.0956-7976.2002.00460.x
- Nilsen, E. S., & Graham, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology*, 58(2), 220–249. doi:10.1016/j.cogpsych.2008.07.002
- Nilsen, E. S., & Graham, S. A. (2012). The development of preschoolers' appreciation of communicative ambiguity. *Child Development*, 83(4), 1400–1415. doi:10.1111/j.1467-8624.2012.01762.x
- Nilsen, E. S., Graham, S. A., Smith, S., & Chambers, C. G. (2008). Preschoolers' sensitivity to referential ambiguity: Evidence for a dissociation between implicit understanding and explicit behavior. *Developmental Science*, 11(4), 556–562. <https://doi.org/10.1111/j.1467-7687.2008.00701.x>
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308(5719), 255–258. doi:10.1126/science.1107621
- Peskin, J., & Ardino, V. (2003). Representing the mental world in children's social behavior: Playing hide-and-seek and keeping a secret. *Social Development*, 12(4), 496–512. doi:10.1111/1467-9507.00245
- Rizzolatti, G., & Fogassi, L. (2014). The mirror mechanism: Recent findings and perspectives. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1644), 20130420. doi:10.1098/rstb.2013.0420
- San Juan, V., Khu, M., & Graham, S. A. (2015). A new perspective on children's communicative perspective taking: When and how do children use perspective inferences to inform their comprehension of spoken language? *Child Development Perspectives*, 9(4), 245–249. doi:10.1111/cdep.12141
- Santesteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., & Bird, G. (2012). Training social cognition: From imitation to Theory of Mind. *Cognition*, 122(2), 228–235. doi:10.1016/j.cognition.2011.11.004
- Scott, R. M., He, Z., Baillargeon, R., & Cummins, D. (2012). False-belief understanding in 2.5-year-olds: Evidence from two novel verbal spontaneous-response tasks. *Developmental Science*, 15(2), 181–193. doi:10.1111/j.1467-7687.2011.01103.x
- Sodian, B., Thoermer, C., & Metz, U. (2007). Now I see it but you don't: 14-month-olds can represent another person's visual perspective. *Developmental Science*, 10(2), 199–204. [10.1111/j.1467-7687.2007.00580.x](https://doi.org/10.1111/j.1467-7687.2007.00580.x)
- Surtees, A. D. R., & Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development*, 83(2), 452–460. doi:10.1111/j.1467-8624.2011.01730.x
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). New York and London: Psychology Press.
- Van Schaik, J. E., & Hunnius, S. (2016). Little chameleons: The development of social mimicry during early childhood. *Journal of Experimental Child Psychology*, 147, 71–81. doi:10.1016/j.jecp.2016.03.003
- Wardlow, L. (2013). Individual differences in speakers' perspective taking: The roles of executive control and working memory. *Psychonomic Bulletin & Review*, 20(4), 766–772. doi:10.3758/s13423-013-0396-1
- Wermelinger, S., Gampe, A., Behr, J., & Daum, M. M. (2018). Interference of action perception on action production increases across the adult life span. *Experimental Brain Research*, 236(2), 577–586. doi:10.1007/s00221-017-5157-3