Training self-other distinction facilitates perspective taking in young children

Kampis, Dora; Duplessy, Helle; Askitis, Dimitris; Southgate, Victoria Helen

Published in:
Child Development

DOI:
10.1111/cdev.13912

Publication date:
2023

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC

Citation for published version (APA):
Perspective-taking entails the adoption of another person's perspective, and this is particularly important in communication. If our perspectives do not align, then in order to interpret someone else's intended referent, we must disregard information known only to us and attend to targets in common ground. It has been claimed that people have an egocentric default which necessitates correction in order to select the other's perspective (e.g., Keysar et al., 2000), a process that is considered effortful and requires executive resources, including inhibitory control (McClery et al., 2011; Qureshi et al., 2010). In adults, several studies have implicated the inferior frontal gyrus, a region broadly involved in inhibition (Aron et al., 2014), in success on tasks that are high in perspective conflict (Hartwright et al., 2015; Samson et al., 2005; van der Meer et al., 2011), suggesting that inhibitory control is required for successful perspective taking.

The demands of perspective selection have long been thought responsible for the relation between children's ability to pass a false belief task and their executive function abilities (Leslie & Polizzi, 1998), as these tasks are argued to require inhibiting one's own prepotent knowledge to report the conflicting (and false) belief of someone else (Leslie et al., 2005). For example, in children, success on a classic test of inhibition, the Dimensional Change Card Sort task (Carlson et al., 1998; Kloo & Perner, 2003), correlates with success on the Sally-Anne location-change false belief task, as do other measures that require inhibition (Devine & Hughes, 2014). There is also a relation between performance on tasks of conflict inhibition, where children need to choose between two conflicting response types, and success on false-belief tasks that are themselves high in perspective conflict (Fizke et al., 2014; Nilsen & Graham, 2009; Sabbagh et al., 2006), suggesting that it is likely the conflict between perspectives that drives this relation. The implication of each of these studies is that the egocentric perspective exerts the dominant influence on behavior and in order to engage in perspective-taking, inhibitory control is required to suppress the self-perspective and select the other's perspective.

Abbreviations: ANCOVA, Analysis of Covariance; DT, Director task; EF, executive function; ToM, Theory of Mind.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. Child Development published by Wiley Periodicals LLC on behalf of Society for Research in Child Development.
Recently, however, work with adults has argued that a separable control process may operate specifically on the management of self and other representations (Santiesteban, White, et al., 2012). In this study, one group of participants took part in an inhibition-of-imitation training prior to completing a perspective-conflict task. Inhibition of imitation, the participant must perform an action opposite from the one they see on the screen, and thus inhibit imitative response tendencies. This task is hypothesized to provide an index of an individual’s ability to enhance the self-representation while inhibiting the other-representation, thus providing a putative measure of self-other control (Brass et al., 2009; Brass & Spengler, 2009; Sowden & Shah, 2014). Following training, participants took part in a communicative perspective-taking task. In this task, a ‘director’ instructs the participant to move objects around a grid. On ‘critical trials’, the director cannot see some objects that the participant can see. To correctly interpret the director’s instruction (e.g., “move the small ball”), the participant must consider not what they themselves can see, but what the director can see (i.e., they should move the smallest ball the director can see, even if this is not the smallest ball that they can see). Participants were reported to look first toward the object that fit the instruction in privileged ground, and only subsequently to the intended referent. This was interpreted as indicating the need to suppress a tendency to respond egocentrically (Keysar et al., 2000). Although work has since pointed to a more nuanced interpretation whereby objects in privileged ground are selected primarily when they are a better perceptual match to the referent and thus participants do not always consider privileged ground first (Hanna et al., 2003; Khu et al., 2020), it remains the case that when the object in privileged ground is a better match for the director’s reference, participants would need to suppress making that choice.

Following training, participants in the imitation-inhibition group were better able to avoid the object in privileged ground and choose the referent that the Director could see, than either a group who were trained to imitate or a control group who were trained with the inhibition of similar actions but outside of a social context. That a motorically equivalent but non-social training did not confer the same advantage was taken as indicating that processes specifically targeting the control of self and other representations are involved in perspective taking (Santiesteban, Banissy, et al., 2012). If indeed a domain-specific self-other control process is involved in perspective taking, above and beyond what is typically considered domain-general cognitive control, it is an important question whether and how such a mechanism contributes to perspective taking in early development. Given that self-representations may emerge slowly (Rochat, 2003), it has been suggested that the need for managing self and other representations may be minimal in infancy but presents a challenge for perspective taking as children reach the preschool years (Southgate, 2020). Evidence suggests that egocentric interference effects are particularly prevalent in young children, but currently we do not know whether a separable cognitive and neural mechanism for handling self-other control could be at play from early in life. Thus, it is important to investigate whether a training that, in adults, results in better perspective taking, ostensibly by affecting the concurrent management of self and other representations, would have a similar facilitating effect at an age where self-representation is still developing, and where it may exert a greater influence on children’s behavior than in adults (Epley et al., 2004).

Nevertheless, young children’s responses on the director task have previously correlated both with their performance on a domain-general stroop-like inhibition task, and on a task involving inhibition of imitation (Nilsen & Graham, 2009). That is, children who are better able to inhibit imitating another’s actions or interfering stimuli were better able to choose the object that is consistent with the director’s perspective, rather than their own. It is thus possible that there is no such distinct self-other control process operating in early development. Various studies suggest that what we identify as different components of executive control in school-age children, rather load on a single factor in the preschool years, implying that early in development, a single construct might underpin performance on a range of tasks that in older children, tap different processes (Gandolfi et al., 2014; Wiebe et al., 2008). Thus, it is possible that distinct control processes only emerge later in childhood, and that initially both social and non-social inhibition training may lead to better performance on the perspective-taking task.

The current study aimed to test whether training preschool children with inhibition of imitation would incur the same benefits for perspective taking as it does in adults. The Director task has previously been adapted for use in several studies in preschoolers (Fan et al., 2015; Nadig & Sedivy, 2002; Nilsen & Graham, 2009). Following Santiesteban and colleagues, three groups of children aged between 3 and 6 years took part in either a short imitation-inhibition training, an imitation training or a non-social inhibition training. Following the training, all children took part in the Director task (Keysar et al., 2000) in which a ‘director’ gave instructions to the child to move objects around a grid. On critical trials, some objects are not visible to the director, and so the child must use their knowledge about their differing perspectives to disambiguate to which objects the director is referring. To establish whether there is a specific role for self-other distinction in successful perspective-taking in early childhood, a control group was given equivalent motor inhibition training experience, but without a self-other context. Furthermore, to further investigate whether any benefit of the critical imitation-inhibition
training is isolated to contexts requiring the participant to manage self and other representations, or may extend to inhibition more broadly, children were tested on two classic inhibition measures: bear-dragon (Reed et al., 1984) and day-night (Gerstadt et al., 1994). Finally, to investigate whether the imitation-inhibition training incurred benefits beyond the Director task to other Theory of Mind components, all children took part in a ToM scale (Wellman & Liu, 2004) at the end of the testing session.

**METHOD**

The present study was preregistered at https://osf.io/kd4hs/?view_only=2d975f33e3bc42be827d3eaa196b188a. Where we substantially diverged from the preregistration, the difference is mentioned.

**Participants**

The present study reports data of $N = 75$ 3- to 6-year-old typically developing Danish-speaking children (36–68 months; mean age = 51 months; SD = 10 months; 42 male; for descriptives per condition see Supplemental Information [SI] 1.1 Participants), tested between 2018 and 2021. Although we originally planned to test 4- to 6-year-olds, we expanded the study's age range to 3 to 6 years to cover the range tested in previous versions of the director task (Nilsen & Graham, 2009). An additional eight children participated but were not included in the final analyses because of technical problems (4), parental interference (1), experimenter error (1), or refusal to complete the study (1). Children were recruited through the national birth registry and tested in the lab ($N = 60$) or recruited and tested on-site in a local kindergarten ($N = 15$) in downtown Copenhagen (Denmark), and were predominantly from white middle-income families, reflective of the local population. Participants received a small gift for participating. The study was approved by the Ethics Committee of the Department of Psychology at the University of Copenhagen, and parents signed an informed consent prior to participation. The participants were assigned to one of three 10-min training conditions: (1) Imitation-inhibition, (2) Imitation, and (3) non-social inhibition, of $n = 25$ each. 10 min was selected because piloting indicated that this was approaching the maximum of what children of this age would tolerate.

**Procedure**

Children were seated in the experimental room with experimenter 1 while experimenter 2 was hidden behind a curtain. Parents stayed in the waiting room, unless the child wanted the parent to be present.

**Training**

Children were assigned to one of three conditions. In each training group, the child was told that practice was over when the bell rang (Figure 1).

---

**FIGURE 1** Examples of training trials in the imitation condition with mouth closed (a) and mouth open (b), and the non-social inhibition condition with flower open (c) and closed (d). In the imitation-inhibition condition children performed the same actions as in the imitation but had to do the opposite action as the experimenter.
In the Imitation-inhibition and Imitation conditions, the experimenter had two identical hand puppets, one for herself and one for the child. The training consisted of opening and closing the hand puppet's mouth. The child was seated opposite experimenter 1 who put the puppet on her hand and asked the child to do the same. The child was told a background story that their puppet was a baby crocodile and for them to become adult crocodiles, they had to learn to open and close their mouths. To practice the movements, the child was first asked to imitate the experimenter, opening the crocodile's mouth when the experimenter opened her crocodile's mouth and closing the crocodile's mouth when the experimenter closed her crocodile's mouth. When the child did the correct movement twice, the experimenter moved on to the training. In the Imitation-inhibition group, the child was instructed to do the opposite of what the experimenter did (i.e., open their crocodile's mouth if the mouth of the experimenter's crocodile was closed, and vice versa), whereas in the Imitation group the child was told to do the same as the experimenter.

As in Santiesteban, Banissy, et al. (2012), we included a non-social imitation-inhibition training group intended to isolate inhibition from self-other distinction so that any effect of the imitation-inhibition training could be attributed to self-other distinction specifically, rather than general executive function resources. In the Non-social inhibition (control) group, the child was seated in front of a screen displaying a photo of a flower-shaped hand puppet on a background of either stars and a moon when the flower was closed or a sun and trees when the flower was open (for details of the protocol see Supporting Information). The flower appeared from the bottom of the screen in an open or closed state. The child was given an identical hand puppet and was told that the flower was a baby flower and for the flower to grow into a big flower, it had to open and close. Like in the Imitation-inhibition group, the child was first asked to show how to open and close the flower and then instructed to do the opposite of the flower on the screen. The training of this group was matched as closely as possible to the imitation-inhibition group, but the flower was not animate and thus the training did not involve the element of doing the opposite of someone else. Experimenter 1 sat next to the child and experimenter 2 was hidden behind the curtain running the script from a computer. To match training between groups, the number of training trials in this non-social inhibition group was yoked to the imitation-inhibition group: every child was randomly assigned to another from the imitation-inhibition group and stopped when they reached the target (maximum) number of training trials of that child, or if they did not reach the maximum number of trials in 10 min, the training stopped. To ensure the yoked design between the imitation-inhibition and non-social inhibition groups, children were randomly assigned to the Imitation-Inhibition or Imitation group which were run concurrently, and the non-social inhibition group was run subsequently.

**Director task**

Following the training, children took part in the Director task. This was implemented based on the description in Fan et al., 2015. The visual stimuli consisted of a 4x4 grid (“shelves”) containing eight different objects (for an example grid see Figure 2). The child sat on one side of the grid, with all objects visible, whereas four
slots were occluded from the view of the director (experimenter 1), who sat on the other side of the grid. The director wore dark sunglasses and maintained her gaze toward the center of the grid when giving instructions, so as not to give the children any indication of which object was to be moved. Three cameras recorded children's behavior. A go-pro camera was placed on the top of the grid to capture children's choice from a close-up front view, another go-pro camera was placed on the child's head on a hat to capture children's head movement and orientation, and a regular video camera recorded the entire scene from behind the child. Some children refused to wear the hat-mounted camera (12) and so their behavior was coded from the other two videos.

First, to ensure that the child understood the game and to show them that from the director's side some objects cannot be seen, the child completed 3 practice trials sitting on the director's side of the grid. Experimenter 2 sat next to the child and helped them give instructions to the director. A first trial with the director moving the right object, was followed by two trials where the director made egocentric errors, moving an object that was occluded from the director's side (now the participant's side). After each error, experimenter 2 asked the child if the object moved by the director was the correct one. If the child said yes, experimenter 2 said “No, you meant this one, didn’t you?”, while pointing at the target object and asking the child to repeat the instruction. The director then moved the correct object.

Following the practice trials, the child and the director switched places. Children listened to auditory instructions from the director who asked them to move precise objects above or under another object. A total of four different grids were presented to the child, one at a time (for all four grids see Supporting Information). For each grid, there were three different-sized objects of the same category, and five other objects each from different categories (see Figure 2). For the three objects of the same category, one of them was occluded from the director's side (e.g., the director could see the largest and medium-sized cars, but not the smallest one).

For each grid, there were three trials: two control trials and one critical trial (out of the three trials of the four grid setups the critical trial was twice the second). On control trials, the child was asked to move an object of which there was only one example, mutually visible to both the child and director (e.g., they were asked to move the banana in the grid in Figure X above the brush). In this case, there was no conflict between the child and the director's perspective. On the critical trials, the child was asked to move the smallest or biggest of the three objects from the same category (e.g., “move the small ball below the apple”). Thus, on critical trials, there was a conflict between the child's perspective and the director's perspective and to respond accurately, the child should ignore the smallest car and move the medium-sized one, which is the smallest car visible to the director. To aid the director, the grids were presented in the same order for all children and the critical object was a car in the first grid, a ball in the second, a glass in the third, and a candle in the last grid. The size of the critical object was also the same for all of the children. For the car and the ball, the child was asked to move the small one, and for the glass and the candle, they were asked to move the big one. The child and director left the room between each grid, while experimenter 2 changed and rearranged the objects for the next grid.

Control tasks

Following the Director task, children participated in two inhibition tasks (Bear-Dragon task and Day-Night) and completed the Theory of Mind (ToM) scale. As previous work has found that Bear-Dragon and Day-Night load on the same factor in children of this age (Reck & Hund, 2011), and were also correlated in our data, (controlled for age, partial r(65) = .263, p = .032), we calculated an inhibition score as the average of the Day-Night and Bear-Dragon performance. If a child did not complete both tasks then we took their performance on the other as their inhibition score.

Bear-Dragon

This task was based on Reed et al. (1984). The child was seated at a table in front of experimenter 1 and presented with two hand puppets: a cat and a dog (used instead of bear and dragon, as the training puppets were crocodiles, potentially resembling a dragon). Experimenter 1 began the task by asking the child to show her how good he/she was at various simple actions, for example, putting their hands on their head, touching their nose, etc., in total 10 different movements (see Supporting Information). The child was then told that the cat was a friendly cat, and that the child should do everything that the cat told him/her to do, but that the dog was not our friend, and that the child should not do whatever the dog asked him/her to do. To make sure the child understood the task instructions, experimenter 1 asked the child to do three different movements in the cat's voice and while wearing the cat on her hand, and the same movements in the dog's voice and while wearing the dog on her other hand. If the child responded correctly, experimenter 1 moved on to the 24 test trials. If, however, the child did not move when asked by the cat or moved when asked by the dog, experimenter 1 explained the task again and asked the child to redo the three different movements, then continued to the test trials when the child was ready. If the child did not understand when explained three times, experimenter 1 still continued to the test trials (this happened in 5 out of 75 children). During the 24 test trials, the child was not corrected, except for the two
first “learning” trials where the experimenter gave feedback, but the child was given the instructions again after the 12th trial independently from their performance.

Day-Night

The Day-Night task was implemented from Gerstadt et al. (1994), and consisted of 16 cards, with a sun printed on half of them and a moon and stars on the other half (see Supporting Information). First, the child was shown a moon card and told that when they saw this card, they should say “day”. Next, they were shown a sun card, and told when they saw this card, they should say “night”. The child was presented with a moon card and asked, “what are you supposed to say?”, followed by a sun card and asked the same question. If the child responded incorrectly in one of these first two cards, these trials were considered practice. In this case, the rules were explained again, and the child was asked about the card with which they made an error. Test trials began if the child responded correctly to the first two cards or if the child still responded incorrectly after being told the rules twice. There was no feedback during any of the test trials.

ToM scale

The ToM Scale was implemented from Wellman and Liu (2004) and consists of seven different stories designed to assess children’s understanding of desires, emotions, knowledge, and beliefs, with incremental difficulty (for the materials used, see Supporting Information). The child was seated across from the experimenter and was shown a series of scenarios with pictures or objects and was asked a set of questions about each scenario. The stories and questions broadly follow an order of difficulty: (i) diverse desires story, (ii) diverse beliefs, (iii) knowledge access, (iv) contents false belief, (v) explicit false belief, (vi) belief-emotion and (vii) real – apparent emotion. The children were never corrected in their answers but received continuous encouragement during each story.

Coding

Training

The training was divided into two halves. The first half was 5 min and the second half was 5 min or less if the child refused to continue, or if the child was assigned to the non-social Inhibition group, where some of the children reached the number of training trials before the 10 min had passed. Scores were calculated for the total number of demonstrations made by the experimenter and the percent of correct responses by the child (how many instances of correct final response occurred during training out of the total). A second coder coded 15 children, 20% of the final sample; inter-rater agreement was high for all measures (see Supporting Information 1.2 Methods).

Director task

For the main measure, we coded how many times children selected the correct object to move in the 4 critical and the 12 control trials. In each trial children received a score of 1 for correct and 0 for incorrect object selected. If a child corrected themselves (first chose an incorrect object but switched to the correct one), we coded it as a correct trial. We also coded number of hesitations (how many times the child touched another object before choosing the final one), whether during a trial they showed looks to distractor and/or the target object prior to making a choice, and based on this we coded their recoveries (there was a look to the distractor but the trial was correct) and incorrect switches (there was a look to the correct target before choosing the distractor). Looks were coded from the two go-pro cameras and the additional video camera (see Procedure – Director task above), based on the direction and center of the camera angle and the child’s head and gaze. A second coder coded 15 children, 20% of the final sample. Inter-rater agreement for correct choice was 100.00%, for hesitation = 90.14% and for (any) looks to distractor and target = 90% and 96% respectively.

Bear-Dragon

The maximum score in the Bear-Dragon was 72 (24 trials × 3). For the bear (cat) trials, we coded 0 = when the child did not move at all, 1 = incorrect movement, 2 = a partial movement (or if the child did something else first and then corrected), 3 = full, correct movement. For the dragon (dog) trials, we coded 0 = when full movement, 1 = incorrect movement, also pointing somewhere or saying “no”, 2 = partial movement or any small involuntary movement and 3 = full inhibition of movement. A second coder coded 15 children, 20% of the final sample; inter-rater agreement was 99.54%.

Day-Night

The maximum score in the Day-Night task was 16. A moon card was coded 1 if the child said “day” and 0 when the child said “night” or anything else. The sun card was coded 1 if the child said “night” and 0 if the child said “day” or anything else. A second coder coded 15 children, 20% of the final sample; inter-rater agreement was 95.30%.
ToM scale

In this task children received a maximum score of 7. Each story was coded 0 = when the child gave the wrong answer (fail) and 1 when the child gave the right answer (pass). Some of the stories had an additional control question which had to be answered correctly to be considered as passed. A second coder coded 15 children, 20% of the final sample; inter-rater agreement was 98.87%.

Confirmatory-exploratory statement

The present paper includes largely confirmatory analyses, the main and critical predictions of which were preregistered and are indicated as such. While the precise type of test used differs at points from the pre-registered, this was done to apply a more fitting analysis, in part due to the review process. The exploratory analysis in the main text is marked, and further ones were included in the Supporting Information (2. Results) for completeness.

RESULTS

All proportion data were arcsine-square-root transformed for analyses, and instead of the preregistered chi-square tests, we used t-test throughout. All p-values are two-tailed. Reported means and plots contain non-transformed values for clarity.

Control trials

We first evaluated whether children in the three conditions were able to follow the director's instructions in the control trials without a distractor. This was not preregistered and served to establish that children had generally understood the task. Accuracy on control trials was high across the three training conditions (imitation-inhibition: $M = 99.5\%$; imitation: $M = 97.5\%$; non-social inhibition: $M = 98.5\%$, see Figure 3b), indicating that the general linguistic and task demands were not too difficult for children of this age. An ANCOVA on the average percentage of control trials on which the target was correctly moved as an outcome, with training (imitation-inhibition, imitation, or non-social inhibition) as a between-subjects factor and age in months as a covariate did not reveal any effect of training ($F(2, 71) = 1.163, p = .318, \eta^2 = .032$), but showed a significant effect of age in months ($F(1, 71) = 5.434, p = .023, \eta^2 = .071$), indicating that children's comprehension of the task did not differ across training conditions, but it increased with age. We have therefore included age as a covariate to the analyses on critical trials.

Main preregistered analyses

Critical trials

As our main analysis, we probed whether training had an effect on children's performance on the critical trials. We analyzed the proportion of trials (out of 4) on which children chose the correct (target) object in the three groups (see Figure 3a). An ANCOVA with training (imitation-inhibition, imitation, non-social inhibition) as a between-subjects factor and age in months as a covariate revealed a significant effect of training ($F(2, 71) = 3.316, p = .042, \eta^2 = .085$), and a significant effect of age ($F(1, 71) = 8.438, p = .005, \eta^2 = .106$) on children's performance on critical trials (for a visualization of performance by condition across the age span, see Supporting Information Figure S2).

Planned and preregistered pairwise comparisons were run to assess which of the groups differed from each other. Specifically, it was predicted and preregistered that the imitation-inhibition group would perform significantly better than the other two. Independent-samples
t-tests on the proportion of critical trials correct (out of 4) showed a significant difference in children's performance between the Imitation-Inhibition group and both control groups (Fisher's LSD; Imitation-inhibition vs. Imitation, t(48) = 2.164, p = .033, 95% CI [0.013 0.303], Cohen's d = 0.62, Imitation-inhibition vs. non-social inhibition, t(48) = 2.26, p = .026, 95% CI [0.021 0.310], Cohen's d = 0.47): children in the Imitation-inhibition group performed better than those in the Imitation and Non-social inhibition groups (M_{imit} = 0.6, SD = 0.279; M_{nonsoc} = 0.41, SD = 0.32; M_{nonsoc} = 0.43, SD = 0.284). There was no significant difference between children's performance on the critical trials in the two control groups (Fisher's LSD; Imitation vs. non-social inhibition, t(48) = 0.996, p = .393, Cohen's d = 0.025). Then, we tested children's performance against chance in each condition. As 19 out of 75 children selected the third object within the category on a critical trial, we considered chance to be 0.33 reflecting a choice between the target, the distractor object, and the third object. Preregistered t-tests against chance on the proportion of children's correct responses in each condition showed that only performance in the Imitation-inhibition group was above chance level (Imitation-inhibition: M = .6, t(24) = 5.164, p < .001, 95% CI [0.37 0.87], Cohen's d = 1.03; Imitation: M = .61, t(24) = 1.809, p = .083, 95% CI [.15 0.26], Cohen's d = 0.36; non-social inhibition: M = .43, t(24) = 1.863, p = .075, 95% CI [.01 0.24], Cohen's d = 0.373).

**Secondary preregistered analyses**

**Recovery**

We analyzed children's looks to the target or distractor object in relation to their subsequent choices. Previously, both adults and children have been reported to make egocentric looks in looking to the distractor object, but adults seem to recover from these looks faster than children (Epley et al., 2004), and bilingual and bilingual-exposed children showed higher rates of recovery than monolingual ones (Fan et al., 2015). This suggests that the ability to correct an egocentric response may be important for reference resolution.

We analyzed whether training had an influence on children's rates of recovery. We reasoned that it would be unlikely that children would remember the location of specific objects and should first need to scan the grid after hearing the instruction. Thus we defined 'recovery' as any look to the distractor object prior to selecting the correct object, on critical trials. An ANCOVA with training (Imitation-inhibition, imitation, or non-social inhibition) as a between-subjects factor and age in months as a covariate revealed a significant effect of training on recovery (F(2, 71) = 4.677, p = .012, η² = .116. In the preregistration we predicted that children in the imitation-inhibition group will be more likely to correct initial egocentric looks. Pairwise comparisons showed a significant difference in amount of recovery between the Imitation-inhibition group and both control groups (Fisher's LSD; Imitation-inhibition vs. Imitation, t(48) = 2.368, p = .021, 95% CI [.028 .333], Cohen's d = 0.706; Imitation-inhibition vs. non-social inhibition, t(48) = 2.868, p = .006, 95% CI [.066 .371], Cohen's d = 0.855): children in the Imitation-inhibition group showed more recovery (M_{imit} = 2.08, SD = 1.152) than the Imitation and Non-social inhibition groups (M_{imit} = 1.32, SD = 1.314; M_{nonsoc} = 1.16, SD = 1.143). There was no difference in recovery between the control groups (Fisher's LSD; Imitation vs. non-social inhibition, t(48) = 0.5, p = .620, 95% CI [.114 .190], Cohen's d = 0.149). We also analyzed children's switches from looks to the target, to choosing the distractor (defined as 'incorrect switch') and found no difference between conditions (see Supporting Information 2.2 Results).

Finally, following the approach of Fan et al. (2015) we compared recovery vs. incorrect switches in the three training groups. Paired-samples t-tests revealed that children in the Imitation-inhibition group showed significantly more recovery than incorrect switches (t(24) = 3.829, p < .001, 95% CI [1.44 4.82]), whereas in the imitation and non-social inhibition groups there was no difference between recovery and incorrect switches (t(24) = 0.296, p = .770, 95% CI [.228 .717]; t(24) = 0.298, p = .769, 95% CI [.212 .758], respectively). Together, these analyses suggest that imitation-inhibition training specifically affected recovery from looks to the distractor rather than a general tendency to choose another object from that at which they had previously looked (for data on children's hesitation in the three conditions see Supporting Information 2.3 Results) (Figure 4).

**Relation to training performance**

We assessed whether children's performance on the training trials (i.e. whether they successfully performed the required action in response to the prompt) influenced their performance on the Director task. As our hypotheses regarding the influence of training concerned the imitation-inhibition training, and in line with this it was the only group who performed above-chance on critical trials and showed an effect on the recovery measure, we focused on this condition only. We analyzed the relation between training performance and both critical trial performance and recovery. To this end, we carried out a median split on the proportion of training trials children responded to correctly (children's average proportion of correct training trials was similar in the three groups: .84 in the imitation-inhibition group, .90 in the imitation group, and .87 in the non-social inhibition group). An ANCOVA with median split based on proportion correct response in training (lower or higher) as between-subjects factors and age in months as a covariate showed...
no significant effect on median split on proportion correct on critical trials, $F(1, 18) = .417, p = .526, \eta^2 = .023$, nor on recovery $F(1, 18) = .142, p = .711, \eta^2 = .008$. This suggests that successfully inhibiting imitation is unlikely to be what drives the relation between training and perspective taking, in turn suggesting that control of self-other representations may not be what training targets.

Exploratory analyses

Relation to training exposure

We reasoned that an alternative means through which training trials could have an effect on critical trial performance is through exposure or highlighting the distinctiveness of self vs. other. Even if children were sometimes unsuccessful at inhibiting imitation, it is possible that the number of training trials they were exposed to benefits critical trial performance by highlighting the difference between self and other. Indeed, children knew that the task was to do the opposite of the experimenter, even if they were not always successful. Furthermore, children were corrected during training if they answered incorrectly. To explore this possibility, we ran a median split analysis based on the number of training trials children were exposed to during the 10-min period. If exposure by itself is important, and may highlight attention to the other's perspective, we could expect the group who received more trials to perform better on the critical trials or show more recovery. An ANCOVA with median split analysis based on the number of training trials (lower or higher) as between-subjects factors and age in months as a covariate showed that the effect on median split on proportion correct on critical trials was $F(1, 18) = 3.982, p = .061, \eta^2 = .181$, and on recovery was $F(1, 18) = 4.498, p = .048, \eta^2 = .2$, see Supporting Information Figure S1.

Additional analyses

Preregistered control tasks

Finally, we asked whether the training conferred a benefit for either domain-general inhibition or Theory of Mind. An ANCOVA with training (Imitation-inhibition, imitation, or non-social inhibition) as a between-subjects factor and age in months as a covariate revealed no statistically significant effect of training on any of the control tasks (ToM scale: $F(2,69) = 0.391, p = .678$; inhibition: $F(2,70) = 1.479, p = .235$), a significant effect of age on the ToM scale and on inhibition (ToM scale: $F(1,69) = 20.02, p < .001, \eta^2 = .225$; inhibition: $F(1,70) = 6.054, p = .016, \eta^2 = .08$), and no other significant effects or interactions.

We further ran an exploratory analysis to ask whether children's performance on the director task was related to their domain-general inhibitory skills. We analyzed whether there was a correlation between the inhibition score and percent of critical trial correct responses or rate of recovery, controlling for age. There was no relation between inhibition score and percent of correct responses, or recovery ($r(66) = .213, p = .081$; and $r(71) = .185, p = .16$, respectively). Furthermore, there was no relation between exposure during training or training performance (across all three training groups) and inhibition score, controlling for age ($r(66) = .101, p = .413$; and $r(66) = .213, p = .081$, respectively).

Path analysis

Using path analysis, we examined the relations between the training, the control tasks (ToM and inhibition), and director task performance (DT), as well as the observed effect of age on DT. We predicted that (1) if training had a significant effect above and beyond any contributions of inhibition or ToM, there should be significant direct paths between the training contrasts and DT, (2) if inhibition or ToM had any effect on DT, we should also observe significant paths between EF, ToM, and DT. Additionally, to investigate the observed effect of age, we probed (3) if age mainly directly affects DT or it has
an effect via inhibition or ToM. We pooled the two control groups (imitation and non-social inhibition) into one control group, as that was the main comparison of interest, and in order to reduce the number of paths to adjust to our sample size [Kline, 2015]. We thus coded the training variable as a contrast between imitation-inhibition and the pooled control group and computed its direct and indirect effects through inhibition and theory of mind.

As can be seen in the model (see Figure 5, details see Supporting Information 2.4 Results), regarding Prediction 1, the direct effect of training on DT (Performance) was significant and accounted for 95% of the total effect. Regarding Prediction 2, we did not observe that either inhibition or ToM have significant paths to DT. Finally, regarding Prediction 3, age had significant direct and no significant indirect effects on DT. The direct effect accounted for 85% of the total effect of age on DT.

**DISCUSSION**

In the current study, we demonstrate that imitation-inhibition confers a similar benefit for perspective taking in children as it does in adults. Three-to-six-year-old children participated in a 10-min training where they either had to imitate, or inhibit imitation of, an experimenter’s actions (imitation and imitation-inhibition group, respectively), or perform an action opposite to the movements of a non-agentive object on a screen (non-social inhibition). Children then participated in the Director task, where on critical trials they had to take the experimenter’s perspective to resolve the referent of her communication, and select the object that she asked them to move. Children in the imitation-inhibition training condition did better on the critical trials where the child and the director’s perspective differed, than children who were in either the imitation training or the non-social inhibition training conditions.

We observed a difference between the imitation-inhibition group and the other two groups, on both the proportion of critical trials that were passed, and in children's ability to recover from looks to the distractor object prior to making a choice. Importantly, children in the imitation-inhibition training group did not simply switch more from any initial gaze but showed a significantly higher tendency for correct than incorrect switches. Children in the other two conditions showed an equivalent number of correct and incorrect switches. However, we found no effect of training on children's performance on the two classic measures of inhibition, nor on the Theory of Mind measures, suggesting that what is tapped by inhibition of imitation is restricted to a demand within the Director task. This was confirmed by the path analysis revealing significant direct effects of training on Director task performance, but not via inhibition or Theory of Mind. In what follows, we consider what may be the best explanation for the beneficial effect of training imitation inhibition on perspective taking.

**Evidence for self-other control?**

Given that shared representations are thought to underlie the tendency for automatic imitation (Brass & Heyes, 2005), training inhibition of imitation is proposed to enhance the control of these shared representations, because the participant has to prioritize the self-relevant action or ignore the other's action (Brass et al., 2009). By doing so, our representations of self and other should gain greater distinction. Imitation-inhibition, through making salient the distinction between self and other and forcing attention to self-relevant information, is thought to have carryover effects to the Director task where control of self and other is vital to prevent interference from self-relevant privileged information when the task should be to prioritize other-relevant information. Thus, just as a more clearly differentiated self and other should render one less susceptible to interference from others’ actions (automatic imitation) in the first place, so too should experience training inhibition-inhibition lead to enhanced separation of self and other. On this assumption, the benefits for social cognition are also assumed to arise from enhanced self-other distinction. Given that an egocentric bias has been documented in the Director task (Epley et al., 2004; Keysar et al., 2000), it is plausible that a reduced tendency to conflate self and other achieved by imitation-inhibition training could result in improved control of self and other representations, reducing the magnitude of a potential egocentric bias.

The original finding of a training effect on perspective taking made a case for distinguishing between
domain-general inhibition and domain-specific self-other control, with self-other control proposed as a process or mechanism integrating or switching between self and other perspectives (Santiesteban, White, et al., 2012). The distinction was made because an equivalent non-social inhibition training did not confer the same benefits for perspective-taking, as we also found was the case for young children in our study. Furthermore, distinct brain regions, more implicated in self-other processes appeared to be commonly involved in both inhibition of imitation and perspective taking (Santiesteban, Banissy, et al., 2012; Santiesteban et al., 2015; Spengler, von Cramon, et al., 2010), supporting the contention that imitation inhibition was targeting a distinct type of social control. In the current study, in addition, we did not observe any effect of imitation inhibition training on measures of general inhibitory control, suggesting that this training did not benefit inhibition more broadly.

The interpretation that imitation-inhibition training is specific to self-other control could seem inconsistent, however, with previous findings showing that domain-general cognitive control, including competence with inhibiting imitation, is implicated in communicative perspective taking. For example, adult's performance on the Stroop task predicts their success at resolving communication in a version of the Director task (Brown-Schmidt, 2009) and 2 to 5-year-old children's looking to the distractor object on a Director task is predicted by their inhibitory control skills (Nilsen & Graham, 2009), although this relation was not observed in simpler versions of communicative perspective-taking tasks involving fewer objects (Brezack et al., 2021; Khu et al., 2020). In our study, we observed no relation between training (either performance or exposure, regardless of group) and our inhibition measures (Day-Night and Bear-Dragon), even though motor inhibition is common to all training groups and the demands of Bear-Dragon. It thus seems plausible that the absence of a relation indicates that—at least for children of this age—the 10-min training was insufficient to confer benefits for this component of inhibitory control.

Nevertheless, the training was sufficient to enhance performance on critical trials of the Director task and so raises the question, if the motor inhibition component was insufficient to enhance performance on this aspect, why it appears sufficient to enhance another dimension of control? Given that a recent meta-analysis casts doubt on the proposal that imitation inhibition recruits specifically ‘social’ brain regions and concludes instead that there is more evidence for recruitment of brain regions traditionally involved in domain-general inhibitory control in imitation inhibition (Darda & Ramsey, 2019), below we consider an alternative explanation for the relation between imitation-inhibition and Director task performance.

If not control, then what?

Previous work has tended to conflate, and/or assume equivalence between self-other control and self-other attention (Bukowski & Samson, 2021). While increased attention to self-other control of self-other representations is what is enhanced. In principle, one could become more aware of self-relevant information, and—in conflict situations more aware of the difference between self and other, without actually becoming any better at managing or switching between these representations. Bukowski and Samson have proposed that there are two core dimensions relevant for successful perspective taking: (i) the relative attention focus on the self vs. other perspective, and (ii) the ability to manage conflicts between the self and other perspective and considering these as separable dimensions had greater predictive power than using a unidimensional measure of perspective-taking (Bukowski & Samson, 2017). Importantly, in recent work with adults, Bukowski and colleagues report that imitation inhibition training was not related to control of self-other representations, but rather to self-salience measures (Bukowski et al., 2021). This data indicates that after imitation-inhibition training, participants are more likely to show an advantage for memory of self-relevant items. To the extent that performance on the director task requires greater attention to the other (not the self), a plausible explanation is that the imitation-inhibition training highlights the distinction between self and other and thereby attention to the other's perspective when considering the target referent.

Previously, a similar distinction was made which argued that perspective-taking entails an initial process of perspective calculation, followed by a process of perspective selection (Leslie & Polizzi, 1998), akin to the notion of self-other control. Neuroimaging evidence has revealed distinct components that may map on to an early perspective calculation stage and a later perspective-selection stage (McCleery et al., 2011) and Qureshi and colleagues have demonstrated that executive functions are required for the latter process of perspective selection, but not for the initial process of perspective calculation (Qureshi et al., 2010). One possibility is thus that imitation-inhibition training does not target the later process of perspective selection akin to the control of self-other representations, but rather the initial process of perspective prioritization (Bukowski & Samson, 2017) or calculation (Leslie & Polizzi, 1998). This would still point to a social mechanism affected by imitation-inhibition training, consistent with the fact that neither in our study nor in Santiesteban et al (2012) was there any effect of non-social training on perspective taking; but this mechanism may involve social attention allocation and sensitivity rather than inhibition or control of self-other perspective.
Self-other attention in development

Findings from developmental work are relevant for this discussion. A recent study investigating the relation between automatic imitation and perspective-taking in 3-year-olds suggested that children of this age may experience a decreased interference from the others' actions. Specifically, Brezack and colleagues reported that only roughly a third of children showed the classic congruency effect where one is slower to respond on imitation-inhibition compared to imitation trials (Brezack et al., 2021). This might suggest that, at this age, children are less susceptible to automatic imitation. While automatic imitation in adults is thought to be due to a failure of control (Spengler, Brass, et al., 2010), a reduced susceptibility to automatic imitation in children of this age - who notoriously have less developed inhibitory control skills than adults—is unlikely to indicate that children do not have a problem with control. Indeed, 3-year-old children can reportedly take another person's perspective but have difficulty flexibly switching between the other's perspective and their own (Moll et al., 2013). Given that this is also the age that children show an egocentric bias on tasks of perspective conflict (Birch & Bloom, 2003), one possibility is that both phenomena are underpinned by a particularly self-focused period of development. Thus, a difference in attention allocation (more self than other-oriented) could explain why children are not as susceptible to automatic imitation.

If children's attention is more self-focused than adults, control may be less relevant for inhibition of imitation than it is for adults. If so, it might be a further hint that the link between imitation inhibition and perspective-taking in the current study is mediated by something other than control of self and other representations. The between-subjects nature of our own design does not allow us to assess whether children show the congruency effect. However, the benefit for perspective taking in the imitation-inhibition training group does suggest that even if children of this age experience less interference from others' actions, highlighting the self-other distinction may still enhance children's perspective-taking abilities by increasing awareness of the difference between self and other.

It is notable that our data suggest that training exposure, rather than performance, was related to Director task performance. We found that the number of trials received during the training (exposure) was positively though not significantly related to critical trial performance, and significantly related to their ability to recover from egocentric looks; but the number of training trials in which they successfully inhibited imitating (performance) was not related to either of these measures. Arguably, enhanced control of self and other representations would result from successful control of imitation on the training. The fact that exposure, not performance, was more relevant is consistent with the interpretation that the training did not improve children's ability to control self and other representations, but rather highlighted or made salient, the differences between self and other, or served as a conceptual reminder that the self and other are different. In line with this interpretation, Fan and colleagues reported that performance on the Director task is better not only in bilingual-speaking children but also in bilingual-exposed children, compared to a monolingual non-exposed group (Fan et al., 2015). While bilingual children have been alleged to have better executive functions due to the requirements of shifting between languages (Kovács & Mehler, 2009), bilingual exposed children are just aware that people speak different languages. The absence of difference between bilingual and bilingual-exposed children may similarly suggest that it is awareness of differences between self and other that facilitates perspective taking on the Director task.

While the imitation inhibition training facilitated children's ability to consider the other's perspective in online communication, possibly by highlighting the distinctness of self and other, many questions remain. While improving perspective taking in this task, the benefits of training did not extend to improvements in children's ability to answer questions about Theory of Mind. While there are likely many differences in the requirements of tasks in which the child is an actor rather than an observer (Moll et al., 2022), it is important to keep in mind that even in adults, there may be little convergence in individual differences between two tasks which both entail the need to separate self and other (Qureshi et al., 2020).

Concluding remarks

Together, in the present study we found that a brief, 10-min imitation-inhibition training enhanced 3- to 6-year-old children's performance on a communicative perspective taking task, compared to a group who received an imitation training. A similar effect was not observed in a group who received a non-social inhibition training, pointing to a specifically social effect. We propose that the imitation-inhibition training enhances children's awareness that self and other are distinct and have unique angles on the external world, ultimately leading them to allocate more attention to the other's perspective when interpreting their communication.

ACKNOWLEDGMENTS

The authors thank N. Husen, L. Berthelsen, P. Steinmeier, A-S. Riise, C. Bering, A. Jeppesen, and K. Habdank for assisting with data collection, and Steven Burton for statistical consultation. D.K., H.L.D, and V.S. are supported by a Consolidator Grant from the ERC (DEVOMIND 726114) to V.S. The data necessary to reproduce the analyses presented here are publicly accessible at: https://osf.io/v8jfp/. The analytic code necessary to reproduce the analyses presented in this paper is not
publicly accessible. The materials necessary to attempt to replicate the findings presented here are publicly accessible in the Supplemental Information and at: https://osf.io/v85jp/ The analyses presented here were preregistered, unless otherwise noted in the manuscript. The preregistration is available at: https://doi.org/10.17605/OSF.IO/KD4HS.

FUNDING INFORMATION
This study was funded by grants from the FP7 Ideas: European Research Council (No. DEVOMIND 726114).

ORCID
Dora Kampis https://orcid.org/0000-0003-1678-3874

REFERENCES
SELF-OTHER DISTINCTION ENHANCES PERSPECTIVE TAKING


SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.